

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:27:07 ; Search time 96.19 Seconds
(without alignments)
8107.378 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcgaccacgcgtcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues
Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	68.4	2.2	75	1 T99072	Alpha epithelial s
2	59	1.9	231	1 T21330	Human gene signatu
3	56.2	1.8	795	1 V55830	FLGA insert stabl
4	52.2	1.7	799	1 V55831	Nucleotide sequenc
5	52.2	1.7	9600	1 V21683	Vector plasmid PCM
6	52.2	1.7	10596	1 Q51731	Plasmid pCISEBON f
7	52.2	1.7	10596	1 T40348	Plasmid pCISEBON f
8	52.2	1.7	10596	1 X15650	Nucleotide sequenc
9	50	1.6	1620	1 V60840	Partial human acid
10	48.2	1.5	558	1 T64563	Mouse thymus and a
11	46.8	1.5	1525	1 V22704	Mouse recombinase
12	45.4	1.5	1129	1 X27340	Human secreted pro
13	45	1.4	45	1 T99077	Probe Ena-A7 for a
14	45	1.4	45	1 T99078	Probe Ena-A9 for a
15	44.6	1.4	1096	1 X22275	Human secreted pro
16	44.6	1.4	1631	1 T72795	Beta-1 integrin mo
17	44.6	1.4	6644	1 X33181	Base sequence of t
18	44.6	1.4	7372	1 X33182	Base sequence of t
19	44.6	1.4	7797	1 X33180	Cowpox virus bsr f
20	44.6	1.4	7996	1 X33184	Base sequence of t
21	44.4	1.4	960	1 X28162	Rat Acid sensitive
22	44.4	1.4	3562	1 V60839	Rat acid sensing i
23	44.2	1.4	1632	1 V68059	Neurodegenerative
24	44.2	1.4	1977	1 X27383	Human secreted pro
25	44.2	1.4	2711	1 V68056	Neurodegenerative
26	44.2	1.4	2955	1 V68057	Neurodegenerative
27	44	1.4	1348	1 X24749	Human interleukin-
28	43.8	1.4	2823	1 T35233	Natural killer lyl
29	43.8	1.4	3350	1 T18200	Infectious bursal
30	43.8	1.4	5503	1 V18187	Fancconi anaemia of
31	43.6	1.4	1174	1 T39050	cDNA encoding cell
32	43.6	1.4	1174	1 V39096	Monocomponent endo
33	43.6	1.4	1537	1 V59804	Human secreted pro

34	43.4	1.4	437	1 X22245	Human secreted pro
35	43.4	1.4	1338	1 O65607	Rabbit zona pelluc
36	43.4	1.4	1338	1 V64789	Rabbit 2PC cDNA. I
37	43.4	1.4	1829	1 V26298	Phosphate starvati
38	43.4	1.4	2132	1 X33816	Coding sequence fo
39	43.4	1.4	2549	1 N91467	Sequence of human
40	43.4	1.4	3557	1 N60801	Human pro-growth h
41	43.4	1.4	3784	1 Q05326	Sequence encoding
42	43.4	1.4	5018	1 N70646	Clone 23B6p102 enc
43	43.4	1.4	5018	1 N50150	Sequence of the cD
44	43.2	1.4	1355	1 N80457	Sequence of cDNA e
45	43.2	1.4	2962	1 X28161	Rat Acid sensitive

ALIGNMENTS

RESULT 1	
T99072	ID T99072 standard; DNA; 75 BP.
AC T99072:	AC T99072: (first entry)
DE 24-MAR-1998	DE Alpha epithelial sodium channel a coding sequence.
KW Alpha epithelial sodium channel; alphaENACa; alphaENACb; binding assay;	
KW amiloride-sensitive salt channel alpha subunit; membrane-transport;	
KW salt substitute; salty taste blocker; ss.	
OS Rattus rattus.	
PN US5693756-A.	
PD 02-DEC-1997.	
PF 23-JAN-1995; 376362.	
PR 23-JAN-1995; US-376362.	
PR 28-FEB-1994; US-202654.	
PA (UYJO) UNIV JOHNS HOPKINS.	
PI Blackshaw S, Li X, Snyder SH;	
DR WPI; 98-031814/03.	
DR P-PSDB; W34529.	
PT Alternatively spliced epithelial sodium channel alpha subunit	
PT proteins - useful in screening assays for salty taste enhancers or	
PT blockers	
PS Disclosure; Column 21-22; 33pp; English.	
CC This sequence represents the coding sequence for the alpha epithelial	
CC sodium channel a (alphaENACa). AlphaENACa and alphaENACb (see W34530)	
CC represent the sequences of the invention. The two sodium channels are	
CC alternatively spliced forms of the amiloride-sensitive salt channel alpha	
CC subunit and can be used in membrane-transport or binding assays to	
CC identify substances that enhance or block perception of a salty taste.	
CC Enhancers could be used as salt substitutes and blockers could be used to	
CC mask salty tastes in foods and pharmaceuticals.	
SQ Sequence 75 BP; 11 A; 22 C; 22 G; 20 T;	
Query Match 2.2%; Score 68.4; DB 1; Length 75;	
Best Local Similarity 99.6%; Pred. No. 4.46e-07;	
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 1545 gggcgcttccttgagacagcctggctgtttctccaagtgtcggaagccttagtgt 1604	
Db 6 GGGCGCCTTCCTCTGGACAGCCCTGGGCTGTCTCTCCAAGTGTGCGAAGCCTTAGTGT 65	
QY 1605 gatcaactac 1614	
Db 66 GATCAACTAC 75	
RESULT 2	
T21330	ID T21330 standard; cDNA to mRNA; 231 BP.
AC T21330:	AC T21330: (first entry)
DE 16-AUG-1996	DE Human gene signature HUMGS02692.
KW Human gene signature; messenger RNA; mRNA; relative abundance; frequency;	
KW human; cloning; mapping; non-biased library; diagnosis; detection;	
KW cell typing; abnormal cell function; ss.	

```
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBU) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 859; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 231 BP; 44 A; 57 C; 52 G; 67 T;

Query Match 1.98; Score 59; DB 1; Length 231;
Best Local Similarity 60.9%; Pred. No. 9.5e-05;
Matches 103; Conservative 0; Mismatches 65; Indels 1; Gaps 1;

QY 2925 tgattattgtcacattctttgtctgtccctccctgggtaaactgagctcctgtgt 2984
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 TGCACCATCTTCTTGACATTTGTCTCCCTTCCCACTAGACTGTAAGTGCCTTGC 122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2985 ggtcaggga-tgagattgtctgtttgtatctctccgtctagccagctcccaactt 3043
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 GGTGAGGACTGAATCTTGCCGTTTATGTATGCTCCATGNNNGNCCATCATCTGCTT 182
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3044 ggaagcgttagcaggctactcaataaaggctgttccatcaaaaaaaa 3092
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 GGNCAAGTAGGCAGGNGCTCAATAAATGTTGTNGCATGNNGGNAAA 231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
V55830/c
ID V55830 standard; DNA; 795 BP.
AC V55830;
DE 18-NOV-1998 (first entry)
DE FLGA insert stabilising polypeptide encoding DNA.
DE Fusion protein; stabilising polypeptide; proteolytic degradation;
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; ss.
OS Epstein-Barr virus.
FH Key Location/Qualifiers
FT CDS 1..788
FT /*tag= a
FT /product= "stabilising polypeptide"

WO9822577-A1.
PD 28-MAY-1998.
PF 17-NOV-1997; IB1508.
PR 25-JUN-1997; US-048945.
PR 15-NOV-1996; US-030986.
PA (MASU) MASUCCI M G.
PI Masucci MG;
DR WPI: 98-312463/27.

DR P-PSDB; W79128.
PT New fusion proteins resistant to proteolytic degradation -
PT comprising a core protein with a stabilising polypeptide comprising
PT a peptide sequence containing glycine repeats
PS Disclosure; Fig 3; 120pp; English.
CC This DNA encodes a stabilising polypeptide and is the FLGA insert of the
CC invention. The invention provides a method for increasing the resistance
CC of a core protein to proteolytic degradation that comprises linking or
CC inserting onto or into the core protein a stabilising polypeptide of
CC formula [(Glya)(Glyb)(Glyc)2]n where Glya, Glyb, Glyc are 1-6
CC sequential Gly residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met,
CC Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not
CC be identical from n repeat to n repeat. Alternatively a nucleic acid
CC encoding the stabilising polypeptide can be linked onto or inserted into
CC a nucleic acid encoding a core protein. The fusion proteins of the
CC invention are more resistant to degradation by proteases and, thus, have
CC a longer half-life than the unfused core protein. The products can be
CC used for treating autoimmune diseases, cancer and inflammation. In
CC particular, the core protein may be an IkappaB regulator protein for the
CC treatment of inflammatory bowel disease, or a nitroreductase protein
CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer
CC or other pathological conditions. The fusion proteins can also be used in
CC diagnostic methods such as in vivo imaging.
SQ Sequence 795 BP; 200 A; 104 C; 478 G; 13 T;

Query Match 1.88; Score 56.2; DB 1; Length 795;
Best Local Similarity 51.5%; Pred. No. 0.00063;
Matches 156; Conservative 0; Mismatches 143; Indels 4; Gaps 1;

QY 1875 cgtcatcttgacctctggtcatcacacttctcatgtctagcgggttccgg---a 1930
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 764 CGGCTCAGCAGCTCTCTTCAGACCCGGGCTCTCACTCTCTCTCCGCCCCCTCTCTCT 705
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1931 gccggtactgtctccagagcaggggccagggtgcccaggaggtgacctccactccag 1990
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 704 GCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 645
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1991 attctctctccgctccgcttctgtctctaccctaccataccccaccactctttggccc 2050
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 644 CTCTGTGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 585
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2051 agcaggcgtatgacctccctccctgacctgacagccctccacctgctctatgctactctag 2110
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 CTCTGTGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 525
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2111 gccccagtgcctccactgacctgtgcggcctgactgtttctgctgtgacctgtgcctg 2170
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 CTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2171 cgc 2173
|||
Db 464 CTC 462

RESULT 4
V55831/c
ID V55831 standard; DNA; 799 BP.
AC V55831;
DE 18-NOV-1998 (first entry)
DE Nucleotide sequence of the stabilising sequence-encoding insert.
DE Fusion protein; stabilising polypeptide; proteolytic degradation;
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; ss.
OS Epstein-Barr virus.
FH Key Location/Qualifiers
FT CDS 1..788
FT /*tag= a
FT /product= "stabilising polypeptide"

WO9822577-A1.
PD 28-MAY-1998.
PF 17-NOV-1997; IB1508.
PR 25-JUN-1997; US-048945.
PR 15-NOV-1996; US-030986.
PA (MASU) MASUCCI M G.
PI Masucci MG;
DR WPI: 98-312463/27.
```



```
PF 07-JUL-1998; U13684..
PR 12-SEP-1997; US-058785.
PR 08-JUL-1997; US-051916.
PR 08-JUL-1997; US-051916.
PR 08-JUL-1997; US-051918.
PR 08-JUL-1997; US-051919.
PR 08-JUL-1997; US-051920.
PR 08-JUL-1997; US-051925.
PR 08-JUL-1997; US-051926.
PR 08-JUL-1997; US-051928.
PR 08-JUL-1997; US-051929.
PR 08-JUL-1997; US-051930.
PR 08-JUL-1997; US-051931.
PR 08-JUL-1997; US-051932.
PR 08-JUL-1997; US-052732.
PR 08-JUL-1997; US-052733.
PR 08-JUL-1997; US-052793.
PR 08-JUL-1997; US-052795.
PR 08-JUL-1997; US-052803.
PR 18-AUG-1997; US-055684.
PR 18-AUG-1997; US-055684.
PR 18-AUG-1997; US-055722.
PR 18-AUG-1997; US-055723.
PR 18-AUG-1997; US-055947.
PR 18-AUG-1997; US-055948.
PR 18-AUG-1997; US-055949.
PR 18-AUG-1997; US-055950.
PR 18-AUG-1997; US-055953.
PR 18-AUG-1997; US-055954.
PR 18-AUG-1997; US-055964.
PR 18-AUG-1997; US-055984.
PR 18-AUG-1997; US-056360.
PR 12-SEP-1997; US-058660.
PR 12-SEP-1997; US-058661.
PR 12-SEP-1997; US-058664.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA,
PI Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Zeng Z;
PI WPI; 99-120770/10.
DR P-PSDB; Y02679.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1; Page 262; 464pp; English.
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number, and the clone it is derived from, are
CC detailed in the descriptor line. The gene can be used to generate fusion
CC proteins by linking to the gene to a human immunoglobulin Fc portion
CC (e.g. X27302) for increasing the stability of the fused protein as
CC compared to the human protein only.
CC The invention relates to 123 novel genes and their fragments (nucleic
CC acid sequences: X27311-X27449; amino acid sequences Y02650-Y02780) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 123 polynucleotides, based on
CC which tissues they are most highly expressed in (see X27311 for described
CC uses).
SQ Sequence 1129 BP; 379 A; 194 C; 192 G; 363 T;

Query Match 1.5%; Score 45.4; DB 1; Length 1129;
Best Local Similarity 58.1%; Pred. No. 0.21;
Matches 79; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 2982 tgtggtcaggatgagatgtctgtgttgcctatcccggtcagccagctctccac 3041
DB 950 TTTTAAAGGTATATTTTCCCAATATTTATTTCTGACCTTAACACACTTTCTACT 1009
QY 3042 ttgggacgggtagcagggtactcaataaaggctgttctccatcaaaaaaa 3101
DB 1010 AAAAAATGGTGAGCAATGAAGACAATAAATTTTTCATTTTTCNAAAAA 1069
QY 3102 aaaaaa 3117
```

```
DB 1070 AAAAAAAAAAAAAA 1085
RESULT 13
T99077/c
ID T99077 standard; DNA; 45 BP.
AC T99077;
DT 24-MAR-1998 (first entry)
DE Probe Ena-A7 for alphaENAC coding sequence.
KW Alpha epithelial sodium channel; alphaENACa; alphaENACb; binding assay;
KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
KW salt substitute; salty taste blocker; probe; ss.
OS Synthetic.
OS Rattus rattus.
PN US5693756-A.
PD 02-DEC-1997. 376362.
PF 23-JAN-1995; US-376362.
PR 23-JAN-1995; US-376362.
PR 28-FEB-1994; US-202654.
PA (UYJO ) UNIV JOHNS HOPKINS.
PI Blackshaw S, Li X, Snyder SH;
PI WPI; 98-031814/03.
DR Alternatively spliced epithelial sodium channel alpha subunit
PT proteins - useful in screening assays for salty taste enhancers or
PT blockers
PS Disclosure; Column 6; 33pp; English.
CC This sequence represents a probe for the coding sequence for the alpha
CC epithelial sodium channel a (alphaENACa). AlphaENACa (see W34529) and
CC alphaENACb (see W34530) represent the sequences of the invention. The two
CC sodium channels are alternatively spliced forms of the
CC amiloride-sensitive salt channel alpha subunit and can be used in
CC membrane-transport or binding assays to identify substances that enhance
CC or block perception of a salty taste. Enhancers could be used as salt
CC substitutes and blockers could be used to mask salty tastes in foods and
CC pharmaceuticals.
SQ Sequence 45 BP; 10 A; 10 C; 12 G; 13 T;

Query Match 1.4%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1134 caacaatggtttgctcctgacatgcgcacagacagaatgacctt 1178
DB 45 CAACAATGTTTGTCCCTGACATGTCGCACAGACAGATGACTT 1
RESULT 14
T99078/c
ID T99078 standard; DNA; 45 BP.
AC T99078;
DT 24-MAR-1998 (first entry)
DE Probe Ena-A9 for alphaENAC coding sequence.
KW Alpha epithelial sodium channel; alphaENACa; alphaENACb; binding assay;
KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
KW salt substitute; salty taste blocker; probe; ss.
OS Synthetic.
OS Rattus rattus.
PN US5693756-A.
PD 02-DEC-1997. 376362.
PF 23-JAN-1995; US-376362.
PR 23-JAN-1995; US-376362.
PR 28-FEB-1994; US-202654.
PA (UYJO ) UNIV JOHNS HOPKINS.
PI Blackshaw S, Li X, Snyder SH;
PI WPI; 98-031814/03.
DR Alternatively spliced epithelial sodium channel alpha subunit
PT proteins - useful in screening assays for salty taste enhancers or
PT blockers
PS Disclosure; Column 7; 33pp; English.
CC This sequence represents a probe for the coding sequence for the alpha
CC epithelial sodium channel a (alphaENACa). AlphaENACa (see W34529) and
```

CC alphaENACb (see W34530) represent the sequences of the invention. The two
 CC sodium channels are alternatively spliced forms of the
 CC aniloride-sensitive salt channel alpha subunit and can be used in
 CC membrane-transport or binding assays to identify substances that enhance
 CC or block perception of a salty taste. Enhancers could be used as salt
 CC substitutes and blockers could be used to mask salty tastes in foods and
 CC pharmaceuticals.
 SQ Sequence 45 BP; 11 A; 11 C; 10 G; 13 T;

Query Match 1.4%; Score 45; DB 1; Length 45;
 Best Local Similarity 100.0%; Pred. No. 0.086;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1498 gactaccgaagcagagctctggtggtattgctattataaactg 1542

DB 45 GACTACCGAAGCAGAGCTCTGGGGCTATTGCTATTATAAATG 1

RESULT 15

ID X22275 standard; DNA; 1096 BP.

AC X22275; 18-MAY-1999 (first entry)

DE Human secreted protein gene 20 clone HGCAB62.

KW Human; secreted protein; gene therapy; protein therapy; cancer; weight;

KW tumour; chromosome mapping; forensic; haematological disease; allergy;

KW inflammation; cell proliferation; viral infection; wound healing;

KW modulation; appetite; behaviour; food additive; preservative; ss.

OS Homo sapiens.

PN WO9903990-A1.

PD 28-JAN-1999.

PF 15-JUL-1998; U14613.

PR 18-AUG-1997; US-056361.

PR 16-JUL-1997; US-052661.

PR 16-JUL-1997; US-052870.

PR 16-JUL-1997; US-052871.

PR 16-JUL-1997; US-052872.

PR 16-JUL-1997; US-052873.

PR 16-JUL-1997; US-052874.

PR 16-JUL-1997; US-052875.

PR 22-JUL-1997; US-053440.

PR 22-JUL-1997; US-053441.

PR 22-JUL-1997; US-053442.

PR 18-AUG-1997; US-055683.

PR 18-AUG-1997; US-055724.

PR 18-AUG-1997; US-055725.

PR 18-AUG-1997; US-055726.

PR 18-AUG-1997; US-055946.

PR 18-AUG-1997; US-055952.

PR 18-AUG-1997; US-055985.

PR 18-AUG-1997; US-055989.

PR 18-AUG-1997; US-056359.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Duan R, Feng P, Ferrie AM, Florence KA, Fouad J,

PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE,

PI Yu G.

DR WPI; 99-132234/11.

DR P-PSDB; Y01447.

PT New nucleic acids encoding secreted human proteins - potentially

PT useful for treating and diagnosing diseases and identifying specific

PT binding agents

PS Claim 4; Page 201-202; 251pp; English.

CC The invention relates to nucleic acid sequences (X22211 to X22282)

CC encoding human secreted proteins (Y01383 to Y01454). The secreted protein

CC gene sequences are deposited with the ATCC under deposit number ATCC

CC 209138, 209139 or 209141. Host cells containing vectors comprising the

CC nucleic acid sequences are used for the recombinant expression of the

CC secreted proteins. The polynucleotide and amino acid sequences are useful

CC for preventing, treating or ameliorating medical conditions e.g. by

CC protein or gene therapy. Pathological conditions can be also diagnosed by

CC determining the amount of the new polypeptides in a sample or by the

CC presence of mutations in the new polynucleotides. The nucleic acid

CC sequences, or its fragments, are useful for chromosome identification and
 CC mapping; as antisense and triplex-forming therapeutics; in gene therapy;
 CC for (forensic) identification of individuals; as molecular weight
 CC markers; to identify related sequences or specific mRNA; in preparation
 CC of oligomers and to raise anti-DNA antibodies. Antibodies are useful as
 CC immunoassay reagents (including for in vivo imaging) and therapeutically
 CC to inhibit or activate particular polypeptides. A very wide range of
 CC disorders may be treated with the polynucleotide and polypeptide
 CC sequences, e.g. autoimmune or haematological diseases, allergy,
 CC inflammation, cancer or other forms of cell proliferation, viral or other
 CC infections. The sequences may also be useful in wound healing, to
 CC modulate differentiation of embryonic stem cells, to modulate weight,
 CC appetite, behaviour etc. and as food additive or preservative. The
 CC present sequence represents a gene encoding a human secreted protein
 CC (see descriptor line for gene number and clone identification).
 SQ Sequence 1096 BP; 191 A; 367 C; 303 G; 231 T;

Query Match 1.4%; Score 44.6; DB 1; Length 1096;

Best Local Similarity 66.3%; Pred. No. 0.32;

Matches 59; Conservative 3; Mismatches 27; Indels 0; Gaps 0;

QY 3029 cccagtcctcccacttgggacgggtaggcaggtactcaataaaggctgttcacatacaaaa 3088

DB 994 CCCAAGGACCCCAATTCTGCCTCTTCTGGGCAATAATGAGTGCTCTTTTCAGCARMWA 1053

QY 3089 aaaaaaiaaaaaaaaaaaaaaaaaaaaaa 3117

DB 1054 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1082

Search completed: September 10, 2000, 23:03:35

Job time: 5788 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:24:57 ; Search time 1087.86 Seconds
(without alignments)
12635.297 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcgaccacgcgtcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues 10495684
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_gss1:*
94: gb_gss2:*
95: gb_gss3:*
96: gb_gss4:*
97: em_gss1:*
98: em_gss2:*
99: em_gss3:*
100: em_gss4:*
101: gb_gss5:*
102: gb_gss6:*
103: gb_gss7:*
104: gb_gss8:*
105: gb_gss9:*
106: em_gss5:*
107: em_gss6:*
108: em_gss7:*
109: em_gss8:*
110: em_gss9:*
111: em_gss10:*
112: em_gss11:*
113: gb_gss10:*
114: gb_gss11:*
115: em_gss12:*
116: gb_gss12:*


```
117: gb_gss13: *
118: gb_gss14: *
119: gb_gss15: *
120: gb_gss16: *
121: gb_gss17: *
122: gb_gss18: *
123: gb_gss19: *
124: em_gss13: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
C 1	642.4	20.6	657	34	AI013917	AI013917	EST208592
C 2	439	14.1	766	22	AA0206530	AA0206530	Zq58d03.r
C 3	404.8	13.0	554	47	AL0365530	AL036530	DKFp564I
C 4	372.6	12.0	387	32	AA925917	AA925917	UI-R-A1-e
C 5	348	11.2	499	25	AA339350	AA339350	zt78a10.r
C 6	307	9.8	470	70	AW249071	AW249071	2820891.s
C 7	283.8	9.1	343	45	AI891593	AI891593	ui57a09.x
C 8	282.4	9.1	383	71	AW354686	AW354686	36764.MAR
C 9	254	8.1	382	34	AI017422	AI017422	ou96e12.x
C 10	243	7.8	496	45	AI891958	AI891958	ui57a09.y
C 11	222.2	7.1	315	74	AW605213	AW605213	CM4-DT003
C 12	218.4	7.0	349	71	AW325627	AW325627	17814.MAR
C 13	217.4	7.0	548	30	AA710038	AA710038	Vt38g03.r
C 14	197	6.3	358	20	AA035472	AA035472	zk27a05.r
C 15	168.6	5.4	425	63	AW105326	AW105326	xd60f06.x
C 16	164.4	5.3	233	46	AI956720	AI956720	ui73f09.y
C 17	160.4	5.1	442	33	AA980401	AA980401	ua52g11.r
C 18	153.4	4.9	278	48	AW026193	AW026193	AV026193
C 19	147.2	4.7	340	28	AA95839	AA95839	nn05c01.s
C 20	140.8	4.5	227	48	AV028316	AV028316	AV028316
C 21	138.8	4.5	496	32	AA882436	AA882436	vx44g06.r
C 22	136.4	4.4	227	48	AV027974	AV027974	AV027974
C 23	129.6	4.2	391	20	AA035445	AA035445	zk27a05.s
C 24	128.4	4.1	532	46	AI932372	AI932372	wd27f01.x
C 25	122.2	3.9	186	48	AV024371	AV024371	AV024271
C 26	117.8	3.8	164	71	AW379740	AW379740	RC4-HT025
C 27	106.4	3.4	463	23	AA240885	AA240885	mv22e05.r
C 28	96.4	3.1	402	26	AA459197	AA459197	zx88h05.r
C 29	96.2	3.1	601	38	AI377290	AI377290	te65c10.x
C 30	94.4	3.0	282	40	AI466523	AI466523	vx44g06.y
C 31	93.4	3.0	438	28	AA51653	AA51653	nf98a10.s
C 32	87.2	2.8	360	27	AA477114	AA477114	zu37c01.s
C 33	84.8	2.7	442	33	AA980401	AA980401	ua52g11.r
C 34	84.2	2.7	546	26	AA02792	AA02792	zu49b09.s
C 35	83.8	2.7	339	85	H26938	H26938	y164f07.xl
C 36	82.4	2.6	642	74	AW578023	AW578023	RC1-CN001
C 37	81.8	2.6	411	85	H39921	H39921	Y054Q05.r1
C 38	81	2.6	369	31	AA834340	AA834340	Oe45e10.s
C 39	81	2.6	526	70	AW249694	AW249694	2819742.3
C 40	80.6	2.6	431	88	R48145	R48145	yJ66c11.r1
C 41	80.2	2.6	371	43	AI699795	AI699795	tz13a09.x
C 42	80	2.6	342	32	AA865826	AA865826	O446b09.s
C 43	79.2	2.5	309	42	AI675767	AI675767	wd15b09.x
C 44	79.2	2.5	405	70	AW259007	AW259007	um76c03.y
C 45	78.8	2.5	399	29	AA641560	AA641560	nr79d03.s

ALIGNMENTS

RESULT	1
AI013917/c	
LOCUS	AI013917 657 bp mRNA EST 08-JAN-1999
DEFINITION	EST208552 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone RSPB245 3' end, mRNA sequence.


```
FEATURES
  source      Location/Qualifiers
1. .554
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="DFP2p56411062"
  /clone_lib="564 (synonym: hfr2)"
  /tissue_type="brain"
  /dev_stage="fetal"
  /lab_host="Xl-2blue"
  /note="vector: pAMP1; Site_1: NotI; Site_2: SalI"
BASE COUNT  122 a  154 c  141 g  136 t
ORIGIN
Query Match      13.0%; Score 404.8; DB 47; Length 554;
Best Local Similarity 83.2%; Pred. No. 1.3e-90;
Matches 460; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 1358 gcgattgccggtcagaacacctttacccttccaagtatacacagcaggtgtgcattcact 1417
Db 2 GGGATGTTCCGTTCGAGAACCTTTACCTTCAAGTACACACAGCAGGTGTATTCACT 61
QY 1418 cctgttccaggagaaacatgatcaagaagtgtggtgtgcctacatcttaccctaaagc 1477
Db 62 CCTGCTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGTCCCTACATCTTATCCGCGGC 121
QY 1478 ccaagggagtgagttctgtactaccgaagacagagctcctgggctattgtctattata 1537
Db 122 CCCAAGACGTGGAGTACTGTGACTACAGAAAGCACAGTTCCTGGGGGTACTGCTACTATA 181
QY 1538 aactgcagggccttctccttggagacgcctgggctgttctcccaagtgtcggaagcctt 1597
Db 182 ACCTCCAGGTTGACATCTCTCCAGACACCTTGGCTGTTCACCAAGTGGCGGAAGCCAT 241
QY 1598 gtggtgtgatacaactacaactctctgcggctactcaagtggtgacatctgtgaagtccc 1657
Db 242 CGAGCGGACCAAGCTNCCAGCTCTCTGTGTGTACTCAAGTGGCCCTCGGTGACATCCC 301
QY 1658 aggattgatcttcagatgctgtccttgcagaaacttaccattatacaacaaagaa 1717
Db 302 AGGATGGGTCTCCAGATGCTATCGGCACAGAACAAATTACACCGTCAACACAAAGAGA 361
QY 1718 acggattgcaaaagctcaacatcttcttcaaggagctgaactataaaactaatcggagt 1777
Db 362 ATGGAGTGGCCAAAGTCAACATCTTCTCAAGGAGCTGAACATACAAACCAATTCGAGT 421
QY 1778 ctctctgtcagatggtcagctcctctgtccaaactggggcagctggagcctgtggt 1837
Db 422 CTCCTCTGTACGATGGTTCACCTTCTGTCCAAAGCTGGGCGAGCAGTGGAGCCTGTGCT 481
QY 1838 ttggctgcctgcctctctgtgtggagatggcgaagtcattcttcacacctcctggta 1897
Db 482 TCGGCTCCTCGGTGTGTGTGTGTGGAGATGGCTGAGCTCGTCTTTGACCTGCTGGTCA 541
QY 1898 tcacatttctcat 1910
Db 542 TCATGTTCTCAT 554
RESULT 4
AA925917/c
LOCUS
DEFINITION
  AA925917 387 bp mRNA EST 04-JUL-1999
  UI-R-A1-es-9-03-0-UI.s1 UI-R-A1 Rattus norvegicus cDNA clone
  UI-R-A1-es-9-03-0-UI.3' similar to gi|458845|emb|X70497|INENACA R.
  norvegicus mRNA for sodium channel, alpha subunit, mRNA sequence.
ACCESSION
  AA925917
VERSION
  AA925917.1 GI:4237108
KEYWORDS
  EST.
SOURCE
  Norway rat.
  Rattus norvegicus
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
```

1 (bases 1 to 387)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
On Apr 21, 1998 this sequence version replaced gi:3073053.
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-dT track served to identify it as a clone from the normalized
adult kidney library. cDNA Library Preparation: M. Fatima Bonaldo,
Ph.D. Clone distribution: clones will be available through Research
Genetics This clone is also available through the I.M.A.G.E.
Consortium at LNL (info@image.llnl.gov). IMAGE ID=1771734
Seq primer: M13 Forward
PolyA-No.

Location/Qualifiers
1. .387
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-A1-es-9-03-0-UI"
/clone_lib="UI-R-A1"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-A1
library is a subtracted library derived from the UI-R-A0
library. The UI-R-A0 library consisted of a mixture of
individually tagged normalized libraries constructed from
rat placenta, adult lung, brain, liver, kidney, heart,
spleen, ovary, and muscle. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-A1) was constructed as follows: PCR
amplified cDNA inserts from a pool of approximately 3,840
UI-R-A0 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-A0
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was
purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-A1 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996)."

BASE COUNT 131 a 70 c 102 g 84 t
ORIGIN

Query Match 12.0%; Score 372.6; DB 32; Length 387;
Best Local Similarity 97.7%; Pred. No. 1.3e-82;
Matches 378; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2714 tcatttgatacacaccatccccagtcctctgtggtgtgtctctctctctctctctctc 2773
Db 387 TCATTGTGATACACATACACCCAGTGCCTCTGTGGTGTGTCTCTCTCTCTCTCTCTC 328
QY 2774 agcctggagagctcttcttactgtccctcagaacacacagctatgatcccttagaacc 2833
Db 327 AGCTTGGAAAGCTTCTTCACTGTCCCTCAGAACACACACACTTTATGCATCCCTAGAACT 268
QY 2834 tgcacaaaggaccattcttttttgaaactcagttaccctgtattgtctccccagaatt 2893
Db 267 TGCCCAAGAGCCACCTTCTTTTGTGACTCAGTTACCCTGTATGTCTTCTCCCAAGATT 208

Qy 2894 gtgcctctccctccaccacccagtgactgtattatgttccacattctttgtgtc 2953
|||||
Db 207 GTGCCCTTTCCCTCACCCCCACAGGTACTGTATTATATGTTTCACATCTTTTGTGTC 148

Qy 2954 tgcctccctgggtaaacagtcctctgtgtgtcagggatgagattctctgtttgt 3013
|||||
Db 147 TGCTCCTTGGGTAAATGTAGCTCCTTTGTGTGTCAGGGATGAGATTGCTCTGTTTGT 88

Qy 3014 atccttcggctcagccagctctccacttgggacgggttaggcaggtactcaataaaaggc 3073
|||||
Db 87 ATCTCTCCGGTTAGCCCGACTTCCCACTTGGGACGGGTAGGCAGGTACTCAATAAAGGC 28

Qy 3074 ttgtccatcaaaaaaa 3100
|||||
Db 27 TTGTTCCATCAAAAAA 1

RESULT 5
AA393950 499 bp mRNA EST 12-AUG-1997
LOCUS zt78a10.r1 Soares.testis.NHT Homo sapiens cDNA clone IMAGE:728442
DEFINITION 5' similar to gb:L29007.cds1 AMILORIDE-SENSITIVE SODIUM CHANNEL
ALPHA-SUBUNIT (HUMAN);, mRNA sequence.

ACCESSION AA393950
VERSION AA393950.1 GI:2046919
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 499)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wyllie,I., Waterston,R. and Wilson,R.

TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1837 Std Error: 0.00
Seq primer: -28m13 rev2 Et from Amersham
High quality sequence stop: 459.

FEATURES
source
1..499
/organism="Homo sapiens"
/db_xref="GB:5925355"
/db_xref="taxon:9606"
/clone="IMAGE:728442"
/clone_lib="Soares.testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5].
TGTACCATCTGAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
111 a 138 c 127 g 121 t 2 others

Query Match 11.2%; Score 348; DB 25; Length 499;

Best Local Similarity 83.3%; Pred. No. 2e-76;
Matches 408; Conservative 0; Mismatches 80; Indels 2; Gaps 1;

Qy 1427 aggaagaacatgatacaagaagtgtgctgtgctacatcttctaccctaagcccaaggag 1486
|||||
Db 1 AGGAGAGCATGATCAAGGAGTGTGGCTGT--CTACATCTTCTATCCGGGCCCCAGAACG 58

Qy 1487 ttgagtctctgactaccgaagcagagctcctgggctattgtctattataaacctgcagg 1546
|||||
Db 59 TGGAGTACTGTGACTACAGAAGCACAGTTCCTGGGGGTACTGCTACTATAAGCTCCAGG 118

Qy 1547 ggcctctctctggacagcctgggctgtttctccaaagtctcggaagcctttagtgta 1606
|||||
Db 119 TTGACTTCTCTCAGACACCCTGGCTGTTTCACCAAGTCGCGAAGCATCGACGCTGA 178

Qy 1607 tcaactacaaactctctgcggctactcaacggtggccatctgtgaagctccacaggatgga 1666
|||||
Db 179 CCAGCTACCACTCTCTGCTGTTACTCACAGTGGCTCGGTGACATCCCAGGAATGGG 238

Qy 1667 tcttcgagatgctctcttgcagacaattacactatttaacacaaagaacgaggtg 1726
|||||
Db 239 TCTTCCAGATGCTATCGCGACAGAACAAATTAACCGTCAACAAGAGAATAATGGAGTGG 298

Qy 1727 caaagctcaacatcttcttcaagagctgaactataaacataattcggagctctctctt 1786
|||||
Db 299 CCAAGTCAACATCTTCTCAAGGAGCTGAACACTACAAAACCAATTCAGTCTCCTCTG 358

Qy 1787 tcaagatggtcagctcctcttccaaacctggcagccagtgagcctgtgttgctcgt 1846
|||||
Db 359 TCAGATGTTGTCACCTCTCTGTCACCAACCTGGCAGCCAGTGGAGCTGTGGTTCGCTCT 418

Qy 1847 ccgtgctctctgtgtgagatgctgcgagcgtcatcttcgacctcctcgtcatcacctc 1906
|||||
Db 419 CGGTGTTGTCGTGTGTGAGATGGCTGAGCTCGCTTTTGACCTCTCTGCTCATCATGTTCC 478

Qy 1907 tcatgctgct 1916
|||||
Db 479 TCATGCTGCT 488

RESULT 6
AA249071
LOCUS AW249071 475 bp mRNA EST 07-JAN-2000
DEFINITION 2820891.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820891 5',
mRNA sequence.
ACCESSION AW249071
VERSION AW249071.1 GI:6592064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 475)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT On Jul 7, 1999 this sequence version replaced gi:5866355.
Other_ESTs: 2820891.3prime
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu
Plate: LLCMS row: G column: 4

Qy 1738 atctcttcaaggagctgaactataaaactaatcggagtctc 1780
 Db 43 ATCTTTTCAAAAGCTGAACATAATAAACTAAATTCGGAGTCTC 1

RESULT 8

AW354686 383 bp mRNA EST 02-FEB-2000
 LOCUS 36764 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION AW354686
 ACCESSION AW354686
 VERSION AW354686.1 GI:6853676
 KEYWORDS EST.
 SOURCE Bos taurus.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 383)
 Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
 Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and
 Keefe,J.W.
 Design and use of four pooled tissue normalized cDNA libraries for
 EST discovery in cattle

Unpublished (2000)
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACACGCTATGACCAT
 BACKWARD: GTTTCCTCCAGTCACGACG
 Plate: 21 row: K column: 24
 Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1..383
 Location/Qualifiers
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 2BOV"
 /tissue_type="pooled"
 /lab_host="DH10p"
 /note="Vector: PCMV SPORt6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."
 BASE COUNT 86 a 98 c 114 g 85 t
 ORIGIN

Query Match 9.1%; Score 282.4; DB 71; Length 383;
 Best Local Similarity 83.9%; Pred. No. 4.6e-60;
 Matches 319; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 1218 ggtgcatggtcaggatgagctgctttatgatgatgggttcctcaacttgagcctgg 1277
 Db 4 GGTGCATGAACGGGATGAGCCTGCTTTATGGATGATGCTGGCTTTAATCTCGCGCCTGG 63

Qy 1278 cgtggagacctccatcagatgagaaaggagccctggagacccctcgagagaaattacgg 1337
 Db 64 CTGGGAGACCTCCATCAGCATGAGTAGAGGAAGCTGTCGACAGACTTGGGGTGACTATGG 123

Qy 1338 cgactgtactgagaatggtagcgtgctcccggtcgaagaacctttacccttcccaagtatac 1397
 Db 124 CGACTGCACCAAGAATGGCAGCGAGGTCGCCAGTGTGAGAACCTTTACAAATACCAAGTACAC 183

Qy 1398 acagcaggtgtgattcaactcctgctcccgaggagacaatgatacaagaagtgtggtctgc 1457
 Db 184 GCAGCAGGTGTGCATCCATTCGTGCTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGTGC 243

Qy 1458 ctacatcttacccttaagcccaaggaggttgagttctgtactacccgaaagcagagctc 1517
 Db 244 CTACATCTTCTACCGCGCGCGGCTAGAGTTCTGTGACTACAGAGAGCATAAATTC 303

Qy 1518 ctgggggtattgtctattataaaactgcaggcgccctctcctcttgacagccttggcctgttt 1577
 Db 304 CTGGGGTACTGCTACTATAAGCTCCAGGATGCCTCTCTCCTCAGACCGCCTGGGCTGTTT 363

Qy 1578 ctccaagtgtcggagcctt 1597
 Db 364 CACCAAGTGCCTGGAAGCCAT 383

RESULT 9

AI017422/c 382 bp mRNA EST 27-AUG-1998
 LOCUS ou96el2.x1 NCI-CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1635694 3'
 DEFINITION similar to gb:L29007.cdsl ALPILORIDE-SENSITIVE SODIUM CHANNEL
 ALPHA-SUBUNIT (HUMAN);, mRNA sequence.
 ACCESSION AI017422
 VERSION AI017422.1 GI:3231758
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 382)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap;
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 www-bio.lnlnl.gov/bbrp/image/image.html
 Insert Length: 479 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES
 source
 1..382
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1635694"
 /clone_lib="NCI-CGAP_Kid3"
 /lab_host="DH10p"
 /note="Organ: kidney; Vector: pTT73D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer,
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pTT73 vector. mRNA
 source: 2 pooled kidneys. Library went through one round
 of normalization. Library constructed by Bento Soares and
 M. Fatima Bonaldo."

BASE COUNT 87 a 93 c 96 g 106 t
 ORIGIN

Query Match 8.1%; Score 254; DB 34; Length 382;
 Best Local Similarity 88.7%; Pred. No. 5.8e-53;
 Matches 275; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1038 gaattattccaagttccaccaccctacgtacgggaactgtctacactttccaatgacaagaa 1097
 Db 371 GAATTACTCTCACTCCACCACCGGATGATGGAACCTGCTATATCTTTCAATGACAAGAA 312

Qy 1098 caactccaattctctggtgtcctccatgcctggagtcacaactggtttgtccctgcacct 1157

Search completed: September 10, 2000, 22:14:59

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:56:52 ; Search time 2749.04 Seconds
(without alignments)
2023.578 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcgcaccacgcgtcg.....aaaaaaaaaaaaaaaaaaaa 3117

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 972840 seqs, 892348106 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pil.*
- 8: gb_p12.*
- 9: gb_prl.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: em_fun.*
- 17: em_hum1.*
- 18: em_hum2.*
- 19: em_in.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_sy.*
- 29: em_un.*
- 30: em_v1.*
- 31: gb_htg1.*
- 32: gb_htg2.*
- 33: gb_in1.*
- 34: gb_in2.*
- 35: em_ba1.*
- 36: em_ba2.*
- 37: em_hum3.*
- 38: em_hum4.*
- 39: gb_pr4.*
- 40: gb_htg3.*
- 41: gb_htg4.*
- 42: gb_htg5.*
- 43: gb_htg6.*

- 44: gb_htg7.*
- 45: em_htg1.*
- 46: em_htg2.*
- 47: em_htg3.*
- 48: em_hum5.*
- 49: gb_pl3.*
- 50: gb_pr5.*
- 51: gb_htg8.*
- 52: gb_htg9.*
- 53: gb_htg10.*
- 54: gb_htg11.*
- 55: gb_htg12.*
- 56: gb_htg13.*
- 57: gb_htg14.*
- 58: gb_in3.*
- 59: gb_htg15.*
- 60: gb_htg16.*
- 61: gb_htg17.*
- 62: em_htg4.*
- 63: em_htg5.*
- 64: em_htg6.*
- 65: em_htg7.*
- 66: em_hum6.*
- 67: gb_htg18.*
- 68: gb_htg19.*
- 69: gb_htg20.*
- 70: gb_htg21.*
- 71: gb_htg22.*
- 72: gb_htg23.*
- 73: gb_htg24.*
- 74: gb_htg25.*
- 75: gb_htg26.*
- 76: gb_htg27.*
- 77: gb_htg28.*
- 78: gb_htg29.*
- 79: gb_htg30.*
- 80: gb_htg31.*
- 81: gb_v11.*
- 82: gb_v12.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3117	100.0	3117	12	RNENACA	X70497 R. norvegic
2	2796	89.7	3081	12	RNASNAC	X70521 R. norvegic
3	2098	67.3	2200	12	RNU54700	U54700 Rattus norv
4	2047	65.7	2200	12	RNU54699	U54699 Rattus norv
5	283	9.1	749	12	AF082073	AF082073 Rattus no
6	283	9.1	2223	12	AF081783	AF081783 Rattus no
7	141	4.5	252	12	RNCDK3UTR	X89818 R. norvegicu
8	104	3.3	3000	12	AF112185	AF112185 Mus muscu
9	98	3.1	98	5	I78503	I78503 Sequence 14
10	71	2.3	852	12	AF002665	AF002665 Rattus no
11	59	1.9	2695	12	CPQ249296	AJ249296 Cavia por
C 12	45	1.4	45	5	I78500	I78500 Sequence 11
C 13	45	1.4	45	5	I78501	I78501 Sequence 12
14	41	1.3	2915	3	OCU132108	AJ132108 Oryctolag
15	40	1.3	636	8	PAU95179	U95179 Prunus arme
16	40	1.3	1620	4	GGVITRO	Y11030 G.gallus mr
17	40	1.3	2171	9	AB002533	AB002533 Homo sapi
18	40	1.3	61001	74	AC034204	AC034204 Homo sapi
19	40	1.3	80861	73	AC036192	AC036192 Homo sapi
C 20	40	1.3	95236	67	AC026450	AC026450 Homo sapi
C 21	40	1.3	101259	11	HSDJ900E8	AL109623 Human DNA
22	40	1.3	103804	73	AC020928	AC020928 Homo sapi
23	40	1.3	122223	39	AC007880	AC007880 Homo sapi
C 24	40	1.3	128758	72	AC008932	AC008932 Homo sapi

c	25	40	1.3	141762	10	HS198C21	AL034400 Human DNA	98	1.2	1411	8	ATH011639	AJ011639 Arabidops
c	26	40	1.3	166947	73	AC034243	AC034243 Homo sapi	99	1.2	1525	34	AF212926	AF212926 Pseudaleet
	27	40	1.3	167627	78	AC015968	AC015968 Homo sapi	100	1.2	1597	33	DDPDEI	X16057 Dictyosteli
	28	40	1.3	170125	41	AC007465	AC007465 Homo sapi	101	1.2	1853	4	AF205877	AF205877 Gallus ga
	29	40	1.3	170273	32	ALI139412	ALI139412 Homo sapi	102	1.2	1886	7	PPPIRINA	X67051 P. patens mr
	30	40	1.3	170948	72	AC012312	AC012312 Homo sapi	103	1.2	1848	8	AF012862	AF012862 Petroseli
	31	40	1.3	172749	39	AC009731	AC009731 Homo sapi	104	1.2	1987	33	DDIRAP1A	L09750 Dictyosteli
c	32	40	1.3	193632	43	AC022037	AC022037 Homo sapi	105	1.2	2050	7	SOAHRI	X57073 S. olteracea
	33	40	1.3	193731	11	HS08N2	AL031123 Human DNA	106	1.2	2158	5	I17544	I17544 Sequence 1
	34	39	1.3	583	3	BT01B8	X63219 B. taurus CI	107	1.2	2182	12	MMOX1G	Z15103 M. musculus
	35	39	1.3	606	34	AF206698	AF206698 Epiblena	108	1.2	2259	9	AK000421	AK000421 Homo sapi
	36	39	1.3	1845	34	AF123275	AF123275 Dictyoste	109	1.2	2627	25	ATNEMESP	Y11994 Arabidopsis
	37	39	1.3	2334	9	HUMSODIUM	L29007 Human Kidne	110	1.2	2680	9	AK000323	AK000323 Homo sapi
	38	39	1.3	2562	39	HSASNA	AF060913 Homo sapi	111	1.2	2712	12	AF107352	AF107352 Mus muscu
	39	39	1.3	3151	9	HSASNA	X76180 H. sapiens m	112	1.2	3093	34	DDU77065	U77065 Dictyosteli
	40	39	1.3	4028	5	A39975	A39975 Sequence 8	113	1.2	3175	49	AF029256	AF029256 Kosteletz
	41	39	1.3	4032	5	I23502	I23502 Sequence 8	114	1.2	4126	12	AF102578	AF102578 Mus muscu
	42	39	1.3	6313	10	HSZ92981	Z92981 H. sapiens S	115	1.2	4254	11	HSM802077	ALI137682 Homo sapi
	43	39	1.3	100368	10	HS514K20	AC008731 Human DNA	116	1.2	4788	33	CBU60649	U60649 Caenorhabdi
	44	39	1.3	140026	41	AC005840	AC005840 Homo sapi	117	1.2	4898	34	DMU07629	U07629 Drosophila
	45	39	1.3	157227	40	ALI158203	ALI158203 Homo sapi	118	1.2	5115	4	AF219232	AF219232 Gallus ga
	46	39	1.3	172081	10	CNS01DVS	ALI136018 Human chr	119	1.2	6691	33	DDGDT1	AJ000992 Dictyoste
c	47	39	1.3	188488	39	AC006057	AC006057 Homo sapi	120	1.2	9455	34	AE001378	AE001378 Plasmodiu
	48	39	1.3	194763	56	AC008453	AC008453 Homo sapi	121	1.2	12029	34	AE001408	AE001408 Plasmodiu
	49	38	1.2	75	5	I78504	I78504 Sequence 15	c	1.2	26016	8	SPBC25B2	AL031853 S. pombe c
	50	38	1.2	75	12	S785B2	S785B2 alpha ENACA	c	1.2	30871	7	SC8021X	Z49704 S. cerevisia
	51	38	1.2	344	12	MMU52006	U52006 Mus musculu	c	1.2	33613	12	AF060868	AF060868 Mus muscu
	52	38	1.2	409	33	DDIAC762	K02959 Slime mold	c	1.2	35005	10	HS30266	AL031703 Human DNA
	53	38	1.2	729	39	AF187015	AF187015 Homo sapi	c	1.2	41906	11	HSAC000124	AC000124 Human Cos
	54	38	1.2	960	5	E02152	E02152 DNA encodin	c	1.2	41947	50	HSAC000358	AC000358 Human Cos
	55	38	1.2	1576	11	MMU801341	ALI133070 Homo sapi	c	1.2	44560	39	AC004536	AC004536 Homo sapi
	56	38	1.2	2345	12	MMU807620	U87620 Mus musculu	c	1.2	70875	10	HSDJ875H3	AL049733 Human DNA
	57	38	1.2	2951	11	HSM801123	ALI17590 Homo sapi	c	1.2	77835	31	PFMAL13P2_3	Continuation (4 of
	58	38	1.2	3036	11	HSM801281	ALI122118 Homo sapi	c	1.2	97943	39	HSAC002467	AC02467 Human BAC
	59	38	1.2	3152	11	HSM800952	ALI17443 Homo sapi	c	1.2	98845	78	AC004906	AC004906 Homo sapi
	60	38	1.2	3385	12	AF071230	AF071230 Cavia cob	c	1.2	107783	67	AC026445	AC026445 Homo sapi
	61	38	1.2	3586	3	BTU14944	U14944 Bos taurus	c	1.2	108313	32	HSG118G19_3	Continuation (4 of
	62	38	1.2	5078	4	TNU75604	U75604 Tilapia nil	c	1.2	110535	39	AC005091	AC005091 Homo sapi
	63	38	1.2	38705	50	AC000031	AC000031 Homo sapi	c	1.2	120169	39	AC005060	AC005060 Homo sapi
	64	38	1.2	78578	31	PFMAL13P9	AL096783 Plasmodiu	c	1.2	122898	55	AC018366	AC018366 Mus muscu
	65	38	1.2	81562	19	AC004359	AC004359 Drosophil	c	1.2	126088	79	AC047999	AC047999 Homo sapi
	66	38	1.2	137139	10	HSBB455A7	ALI21580 Human DNA	c	1.2	139607	32	HSJ196B13	ALI21968 Homo sapi
	67	38	1.2	140784	54	AC008438	AC008438 Homo sapi	c	1.2	140212	67	AC016790	AC016790 Homo sapi
	68	38	1.2	153009	41	AC009801	AC009801 Homo sapi	c	1.2	148923	68	AC007991	AC007991 Homo sapi
	69	38	1.2	158118	53	AC016537	AC016537 Homo sapi	c	1.2	155236	77	AC009701	AC009701 Homo sapi
c	70	38	1.2	158118	53	AC016537	AC016537 Homo sapi	c	1.2	171198	53	AC021036	AC021036 Homo sapi
c	71	38	1.2	172437	59	AC022270	AC022270 Homo sapi	c	1.2	173508	78	AC004898	AC004898 Homo sapi
	72	38	1.2	174817	43	AC011399	AC011399 Homo sapi	c	1.2	173826	31	AP001160	AP001160 Homo sapi
	73	38	1.2	181693	67	AC026322	AC026322 Homo sapi	c	1.2	174148	73	AC009057	AC009057 Homo sapi
c	74	38	1.2	264851	34	AE003644	AE003644 Drosophil	c	1.2	184234	40	ALI162853	ALI162853 Homo sapi
c	75	38	1.2	30962	34	DROSADH04	AE003410 Drosophil	c	1.2	184292	59	AC016763	AC016763 Homo sapi
c	76	37	1.2	320	13	CNS01NFS	ALI352169 Anopheles	c	1.2	180418	44	AC017095	AC017095 Homo sapi
	77	37	1.2	394	14	SYNHETP505	J02540 heteropolym	c	1.2	180581	60	AC018609	AC018609 Mus muscu
	78	37	1.2	471	33	AF026265	AF026265 Lucilia c	c	1.2	180798	71	AC011351	AC011351 Homo sapi
	79	37	1.2	495	33	S81785	S81785 Myr p II-ma	c	1.2	183558	32	ALI135792	ALI135792 Homo sapi
c	80	37	1.2	530	13	CNS01TLU	ALI45891 Anopheles	c	1.2	184234	40	ALI162853	ALI162853 Homo sapi
	81	37	1.2	631	8	AF044204	AF044204 Gossypium	c	1.2	186908	31	AP001812	AP001812 Homo sapi
	82	37	1.2	650	11	HSM800760	U080231 Homo sapi	c	1.2	190877	31	AP001269	AP001269 Homo sapi
	83	37	1.2	727	8	ATU94495	U94495 Arabidopsis	c	1.2	192581	31	PFMAL13P1	AL049180 Plasmodiu
	84	37	1.2	739	7	ATCYTGPX2	AJ000470 Arabidops	c	1.2	194932	32	CNS01DXH	ALI139316 Homo sapi
	85	37	1.2	739	39	AF067168	AF067168 Homo sapi	c	1.2	195102	60	AC005506	AC005506 Plasmodiu
	86	37	1.2	858	8	OSU74296	U74296 Oryza sativ	c	1.2	204663	56	AC022105	AC022105 Homo sapi
	87	37	1.2	872	33	DMRNAPB	X75499 D. melanogas	c	1.2	204917	39	AC008040	AC008040 Homo sapi
	88	37	1.2	970	5	E00424	E00424 Pig elastas	c	1.2	205804	71	AC027654	AC027654 Mus muscu
	89	37	1.2	996	7	ENASPNDR	Z80229 Emericella	c	1.2	209916	74	AC010422	AC010422 Homo sapi
	90	37	1.2	1001	5	AR016469	AR016469 Sequence	c	1.2	210200	53	AC019238	AC019238 Homo sapi
	91	37	1.2	1012	11	PAU82433	U82433 Prunus arme	c	1.2	214718	79	AC046147	AC046147 Mus muscu
	92	37	1.2	1062	11	HSM802518	ALI62002 Homo sapi	c	1.2	216524	51	AC007305	AC007305 Mus muscu
	93	37	1.2	1107	7	ST27G1433	Y11686 Solanum tub	c	1.2	217304	67	AC010244	AC010244 Homo sapi
	94	37	1.2	1220	7	SOPXR1	Y10462 S. olteracea	c	1.2	220995	69	AC015995	AC015995 Homo sapi
	95	37	1.2	1223	11	HSU933205	U93205 Homo sapien	c	1.2	225490	32	ALI158052	ALI158052 Homo sapi
	96	37	1.2	1341	33	CCLDOPACB	Y11906 C. capitata	c	1.2	230129	41	HSAC76P10	ALI132867 Homo sapi
	97	37	1.2	1381	7	ATPHDIPSN	X83764 A. thaliana	c	1.2	248619	67	AC008875	AC008875 Homo sapi

171	37	1.2	252102	60	AC005140	244	36	1.2	907	8	ATU95035	U95035 Arabidopsis
172	37	1.2	253307	33	PFMAL3P7	245	36	1.2	912	12	RNY17327	Y17327 Rattus norv
173	37	1.2	256172	41	AC005139	246	36	1.2	918	7	ZMNB1B	X66077 Z.mays MNBI
174	37	1.2	282806	60	AC006279	247	36	1.2	925	7	AB010898	AB010898 Daucus ca
175	37	1.2	308930	67	AC016562	248	36	1.2	935	10	S75940	S75940 [Alu repeat
176	36	1.2	62	105	S2152S31	249	36	1.2	946	49	AF010228	AF010228 Lycopersi
177	36	1.2	105	13	HUMUT956B	250	36	1.2	950	82	AF012335	AF012335 Human End
178	36	1.2	212	9	AB023798	251	36	1.2	952	8	AF107026	AF107026 Triticum
179	36	1.2	299	39	AF162270	252	36	1.2	958	4	S54890	S54890 growth horm
180	36	1.2	313	13	G36148	253	36	1.2	961	9	AK000718	AK000718 Homo sapi
181	36	1.2	333	9	HSAL0006	254	36	1.2	972	5	I22272	I22272 Sequence 1
182	36	1.2	341	9	HSC2N4	255	36	1.2	973	7	VRSA	X70671 V.radiata m
183	36	1.2	345	12	RNRPRL38	256	36	1.2	978	4	AF012462	AF012462 Pleuronex
184	36	1.2	346	9	HSALU005	257	36	1.2	979	7	VVCHI	X75963 V.vinifera
185	36	1.2	346	9	HSC4N2	258	36	1.2	982	7	AB024276	AB024276 Citrus un
186	36	1.2	346	11	HSU67810	259	36	1.2	986	39	AF169154	AF169154 Homo sapi
187	36	1.2	361	10	HOMBRC42	260	36	1.2	993	8	PRU90347	PRU90347 Pinus radi
188	36	1.2	380	5	A39972	261	36	1.2	1016	34	TVU87096	TVU87096 Trichomonas
189	36	1.2	380	5	I23499	262	36	1.2	1033	12	MUSUNKN1	MUSUNKN1
190	36	1.2	380	5	I23499	263	36	1.2	1033	12	MUSUNKN1	MUSUNKN1
191	36	1.2	380	11	HSU67832	264	36	1.2	1046	5	A21103	A21103 N.tabacum S
192	36	1.2	397	10	S75201	265	36	1.2	1046	5	I44073	I44073 Sequence 4
193	36	1.2	428	34	AF034637	266	36	1.2	1049	8	AF071889	AF071889 Prunus ar
194	36	1.2	433	5	I93570	267	36	1.2	1050	7	AB029635	AB029635 Daucus ca
195	36	1.2	442	3	OCRPTC	268	36	1.2	1056	12	AF038848	AF038848 Mus muscu
196	36	1.2	462	34	AF155364	269	36	1.2	1061	34	AF001981	AF001981 Dictyoste
197	36	1.2	473	12	RNMDG2	270	36	1.2	1062	7	PALHCB12	PALHCB12
198	36	1.2	490	11	HSU800244	271	36	1.2	1065	3	CHAF000167	CHAF000167 Capra hir
199	36	1.2	500	12	MUSL1BA4	272	36	1.2	1071	11	AF052122	AF052122 Homo sapi
200	36	1.2	503	8	S78730	273	36	1.2	1099	12	AF067790	AF067790 Rattus no
201	36	1.2	514	9	AK000450	274	36	1.2	1103	8	AF020793	AF020793 Medicago
202	36	1.2	514	11	AF015158	275	36	1.2	1105	39	HOMOP2SG1	AF106543 Homo sapi
203	36	1.2	538	13	G15869	276	36	1.2	1110	11	AF014838	AF014838 Homo sapi
204	36	1.2	541	7	SHY16247	277	36	1.2	1113	5	E12888	E12888 cDNA encodi
205	36	1.2	556	49	AF165420	278	36	1.2	1115	10	MACPOMCA	M19658 Macaca neme
206	36	1.2	565	7	TDEHD25	279	36	1.2	1115	11	HSU96074	U96074 Human trans
207	36	1.2	580	11	HSU800144	280	36	1.2	1130	11	HSU802210	AL137484 Homo sapi
208	36	1.2	590	11	HSU67221	281	36	1.2	1136	8	ASU58949	U58949 Allium sati
209	36	1.2	604	33	HAY12285	282	36	1.2	1138	11	AF026807	AF026807 Homo sapi
210	36	1.2	639	11	HSU12285	283	36	1.2	1149	9	AF159804	AF159804 Vigna ung
211	36	1.2	647	7	D88314	284	36	1.2	1160	9	AK000844	AK000844 Homo sapi
212	36	1.2	667	49	AF2339930	285	36	1.2	1162	8	AF117707	AF117707 Lycopersi
213	36	1.2	673	12	RNCEH	286	36	1.2	1173	8	PHRNANAM	X92205 P.hybrida m
214	36	1.2	684	5	AF037084	287	36	1.2	1192	39	AF068235	AF068235 Homo sapi
215	36	1.2	685	7	AB027154	288	36	1.2	1199	39	AF161517	AF161517 Homo sapi
216	36	1.2	701	7	PBL131103	289	36	1.2	1201	49	AF063805	AF063805 Phasodact
217	36	1.2	709	49	AF118113	290	36	1.2	1209	34	AF100287	AF100287 Trypanoso
218	36	1.2	720	34	AF082530	291	36	1.2	1225	7	ZMNB1A	X66076 Z.mays MNBI
219	36	1.2	722	8	AF030252	292	36	1.2	1231	9	AK000476	AK000476 Homo sapi
220	36	1.2	723	7	AB015760	293	36	1.2	1239	12	AF000577	AF000577 Rattus no
221	36	1.2	742	5	I77092	294	36	1.2	1246	7	ATPRXR4GE	X98316 A.thaliana
222	36	1.2	742	7	CACORA16	295	36	1.2	1246	7	PVPFE	X58274 P.vulgaris
223	36	1.2	754	7	RICORFB	296	36	1.2	1247	7	ATFOI	X99096 A.thaliana
224	36	1.2	772	39	MMBGLC2	297	36	1.2	1250	5	A38574	A38574 Sequence 4
225	36	1.2	776	7	AB007819	298	36	1.2	1255	7	AB028149	AB028149 Medicago
226	36	1.2	778	10	HSU46465	299	36	1.2	1259	5	E03348	E03348 cDNA sequen
227	36	1.2	788	33	AF030184	300	36	1.2	1267	11	HSU800196	AL050109 Homo sapi
228	36	1.2	799	11	HSU67233	301	36	1.2	1273	15	AX001570	AX001570 Sequence
229	36	1.2	799	50	AF151080	302	36	1.2	1276	39	AF028823	AF028823 Homo sapi
230	36	1.2	800	7	PCPR2	303	36	1.2	1278	4	AF100931	AF100931 Oncorhyn
231	36	1.2	806	8	CPA011715	304	36	1.2	1281	11	HSU801360	AL133084 Homo sapi
232	36	1.2	808	13	G41948	305	36	1.2	1284	5	A20553	A20553 NF-YB. 6/19
233	36	1.2	825	8	AF052871	306	36	1.2	1298	39	AF113689	AF113689 Homo sapi
234	36	1.2	859	11	HSU802545	307	36	1.2	1303	11	HSU90913	U90913 Human clone
235	36	1.2	869	8	LEU84140	308	36	1.2	1316	3	OCUL131955	AF131955 Oryctolag
236	36	1.2	884	4	PRDB428	309	36	1.2	1319	39	HUMTRNSAL	L19437 Homo sapien
237	36	1.2	887	11	HSU801336	310	36	1.2	1321	11	HSU801113	AL117580 Homo sapi
238	36	1.2	893	5	I52156	311	36	1.2	1321	33	TVU87094	TVU87094 Trichomonas
239	36	1.2	893	5	I52156	312	36	1.2	1329	7	ATH6787	ATJ06787 Arabidops
240	36	1.2	898	11	HSU801961	313	36	1.2	1332	8	OSHOX1	X96681 Oryza sativ
241	36	1.2	900	4	AB029757	314	36	1.2	1332	11	HSU8000047	AL050389 Homo sapi
242	36	1.2	904	34	AF091604	315	36	1.2	1338	1	AR055519	AR055519 Sequence
243	36	1.2	905	4	GRNAPFCP	316	36	1.2	1352	34	OCU67148	U67148 Osterlagia
									1354	11	HSU8001325	AL133054 Homo sapi

317	36	1.2	1362	9	AK000559	AK000559 Homo sapi	390	1.2	1978	11	AF016709	AF016709 Homo sapi
318	36	1.2	1364	7	AB025573	AB025573 Nicotiana	391	1.2	1992	33	AF017442	AF017442 Crassos
319	36	1.2	1366	11	HSB800180	AL050097 Homo sapi	392	1.2	2002	12	MUSBMK	J03023 Murine macr
320	36	1.2	1375	7	EGWWDH	X78900 E.gunnlii mR	393	1.2	2014	39	AF122922	AF122922 Homo sapi
321	36	1.2	1380	8	AF038585	AF038585 Zea mays	394	1.2	2018	9	HSXOC11P	AJ000041 Homo sapi
322	36	1.2	1384	12	RATRC3A	L09119 Rattus norv	395	1.2	2019	11	HSB802272	AL137536 Homo sapi
323	36	1.2	1395	5	CA1575	A41575 Sequence 1	396	1.2	2021	7	ATSHAGGY	Y07822 A.thaliana
324	36	1.2	1395	7	CMGTB	X73314 C.maxima mR	397	1.2	2023	8	AF012899	AF012899 Sambucus
325	36	1.2	1402	11	AF003738	AF003738 Homo sapi	398	1.2	2036	7	PSY17796	Y17796 Pisum sativ
326	36	1.2	1430	12	AF058798	AF058798 Mus muscu	399	1.2	2047	8	AF124369	AF124369 Nicotiana
327	36	1.2	1432	10	IKO75368	AL109672 Homo sapi	400	1.2	2055	8	AGU79997	U79997 Atriplex ga
328	36	1.2	1435	11	HSB801967	AL137298 Homo sapi	401	1.2	2063	11	HSB801958	AL137290 Homo sapi
329	36	1.2	1441	5	A65340	A65340 Sequence 63	402	1.2	2064	39	HSU92544	U92544 Human hepat
330	36	1.2	1442	15	AX001574	AX001574 Sequence	403	1.2	2065	11	HSB801264	AL122107 Homo sapi
331	36	1.2	1449	9	AK000392	AK000392 Homo sapi	404	1.2	2082	5	AR059958	AR059958 Sequence
332	36	1.2	1450	9	HS9112	X57348 H.sapiens m	405	1.2	2095	39	AF147790	AF147790 Homo sapi
333	36	1.2	1451	34	AFZ01906	AFZ01906 Drosophill	406	1.2	2114	5	I22020	I22020 Sequence 6
334	36	1.2	1487	8	AF169800	AF169800 Lillium hy	407	1.2	2126	9	AK000248	AK000248 Homo sapi
335	36	1.2	1488	9	AK000037	AK000037 Homo sapi	408	1.2	2137	9	AK000251	AK000251 Homo sapi
336	36	1.2	1500	4	XLNAKATP	Y11587 X.laevvis mR	409	1.2	2141	8	AF161885	AF161885 Macadamia
337	36	1.2	1503	11	HSB800520	AL050223 Homo sapi	410	1.2	2145	7	CSNADP	X58832 C.sorokinia
338	36	1.2	1504	11	AF052157	AF052157 Homo sapi	411	1.2	2147	12	MSP281AM	M25244 Mouse pre-B
339	36	1.2	1505	12	AFZ15669	AFZ15669 Mus muscu	412	1.2	2160	12	MMNOD	X70514 M. musculus
340	36	1.2	1534	11	HSB801279	AL122116 Homo sapi	413	1.2	2165	12	MMU96685	U96685 Mus musculu
341	36	1.2	1535	39	AF109683	AF109683 Homo sapi	414	1.2	2186	49	AF139205	AF139205 Abies gra
342	36	1.2	1555	11	HSB802203	AL137478 Homo sapi	415	1.2	2189	11	HSB802178	AL137452 Homo sapi
343	36	1.2	1588	5	E03349	E03349 cDNA sequen	416	1.2	2198	12	AB007812	AB007812 Mus muscu
344	36	1.2	1588	11	HSB800676	AL080159 Homo sapi	417	1.2	2201	12	AF109155	AF109155 Mus muscu
345	36	1.2	1590	34	AF132551	AF132551 Drosophill	418	1.2	2202	39	AF126488	AF126488 Homo sapi
346	36	1.2	1596	33	LFRAC21596	X96512 L. terrestris	419	1.2	2220	33	DDCAMP1	X52688 D.discoidleu
347	36	1.2	1602	5	AK000496	AK000496 Sequence	420	1.2	2225	7	OSA243828	AJ243828 Oryza sat
348	36	1.2	1606	39	AF113015	AF113015 Homo sapi	421	1.2	2245	9	AK000482	AK000482 Homo sapi
349	36	1.2	1628	10	HSU08374	U08374 Human cytos	422	1.2	2255	34	AFZ01905	AFZ01905 Drosophill
350	36	1.2	1631	8	U79556	U79556 Arabidopsis	423	1.2	2270	11	HSB802322	AL137574 Homo sapi
351	36	1.2	1632	5	E07941	E07941 cDNA encodi	424	1.2	2271	9	HUMASP	J05032 Human aspar
352	36	1.2	1639	5	AR072149	AR072149 Sequence	425	1.2	2271	11	HSB802304	AL137561 Homo sapi
353	36	1.2	1649	8	ATH7450	AJ007450 Arabidops	426	1.2	2334	7	NEUCAM	M73032 Neurospora
354	36	1.2	1670	11	HSB801364	AL133088 Homo sapi	427	1.2	2337	11	HSB801110	AL117578 Homo sapi
355	36	1.2	1691	9	AK000453	AK000453 Homo sapi	428	1.2	2337	39	AF090886	AF090886 Homo sapi
356	36	1.2	1699	11	HSB801509	AL133640 Homo sapi	429	1.2	2333	5	A92394	A92394 Sequence 19
357	36	1.2	1699	11	HSU39656	U39656 Human MAP k	430	1.2	2333	7	ATSGT	Z83833 Arabidopsi
358	36	1.2	1700	1	BRE237608	AJ237608 Borrelia	431	1.2	2368	34	AF190745	AF190745 Drosophill
359	36	1.2	1704	39	HSU81232	U81232 Human tumor	432	1.2	2381	5	I30339	I30339 Sequence 9
360	36	1.2	1723	49	AF136004	AF136004 Triticum	433	1.2	2381	5	I96063	I96063 Sequence 9
361	36	1.2	1725	8	AF007785	AF007785 Zea mays	434	1.2	2384	5	I30334	I30334 Sequence 1
362	36	1.2	1748	11	AF054988	AF054988 Homo sapi	435	1.2	2384	5	I96058	I96058 Sequence 1
363	36	1.2	1776	7	RCENOLASE	Z28386 R.communis	436	1.2	2394	11	AF061795	AF061795 Homo sapi
364	36	1.2	1792	8	AF170272	AF170272 Manihot e	437	1.2	2409	7	SCYNL142W	Z71418 S.cerevisia
365	36	1.2	1801	12	MMU6584	AJ006584 Mus muscu	438	1.2	2421	10	HSARAL	Y14314 Homo sapien
366	36	1.2	1804	33	ACU89984	U89984 Acanthamoeb	439	1.2	2440	12	RNU39207	U39207 Rattus norv
367	36	1.2	1808	9	HSOMGPA	X51694 H.sapiens O	440	1.2	2445	8	AF019113	AF019113 Oryza sat
368	36	1.2	1817	8	ATJ00163	AJ001363 Arabidops	441	1.2	2472	39	AF151685	AF151685 Homo sapi
369	36	1.2	1838	39	AFZ06502S2	AFZ06503 Homo sapi	442	1.2	2475	11	HSB801926	AF137266 Homo sapi
370	36	1.2	1846	39	AF090940	AF090940 Homo sapi	443	1.2	2486	33	CEU23830	U23830 Caenorhabdi
371	36	1.2	1857	11	AF007142	AF007142 Homo sapi	444	1.2	2499	11	HSB801350	AL133016 Homo sapi
372	36	1.2	1872	9	HSCHRX	X65724 H.sapiens D	445	1.2	2500	34	DDU89350	U89350 Dictyosteli
373	36	1.2	1882	12	RNCALRET	X53363 Rat mRNA fo	446	1.2	2530	11	HSB801333	AL133062 Homo sapi
374	36	1.2	1891	9	HUMENTRALU	L30117 Human heat-	447	1.2	2532	33	SLMACTARD	M15272 Sline mold
375	36	1.2	1893	11	AF070573	AF070573 Homo sapi	448	1.2	2563	7	NEUGPROTAA	L11452 Neurospora
376	36	1.2	1899	12	AF093542	AF093542 Mus muscu	449	1.2	2568	33	DDP8A7	X51947 Dictyostelli
377	36	1.2	1915	11	AF038191	AF038191 Homo sapi	450	1.2	2575	11	HSB801461	AL133653 Homo sapi
378	36	1.2	1929	7	LEGALK	X99851 L.erecta mR	451	1.2	2648	11	HSB801470	AL133656 Homo sapi
379	36	1.2	1930	12	MMU249198	AJ249198 Mus muscu	452	1.2	2675	11	HSB801280	AL122117 Homo sapi
380	36	1.2	1934	4	S59519	S59519 stannioalc	453	1.2	2718	12	AF199027	AF199027 Mus muscu
381	36	1.2	1934	9	AK000074	AK000074 Homo sapi	454	1.2	2727	11	HSB801383	AL133103 Homo sapi
382	36	1.2	1942	9	HSAL010841	AJ010841 Homo sapi	455	1.2	2762	11	AF039916	AF039916 Homo sapi
383	36	1.2	1945	33	CEU61236	U61236 Caenorhabdi	456	1.2	2763	11	HSB802278	AL137541 Homo sapi
384	36	1.2	1952	5	I28450	I28450 Sequence 1	457	1.2	2816	11	HSB800549	AL117440 Homo sapi
385	36	1.2	1952	5	I81145	I81145 Sequence 1	458	1.2	2876	11	HSB800660	AL080147 Homo sapi
386	36	1.2	1952	5	I84698	I84698 Sequence 1	459	1.2	2896	4	AF061981	AF061981 Xenopus 1
387	36	1.2	1957	12	AF098295	AF098295 Mus muscu	460	1.2	2922	11	HSB802222	AL137495 Homo sapi
388	36	1.2	1964	7	SCYNL141W	Z71417 S.cerevisia	461	1.2	2923	7	ATHSP91	Z70314 A.thaliana
389	36	1.2	1968	10	HSU09115	U09115 Human retro	462	1.2	2924	11	HSB800664	AL080150 Homo sapi

463	36	1.2	2924	49	AF188678	AF188678 P1sum sat	c 536	36	1.2	12593	34	AE001418	AE001418 Plasmodiu
464	36	1.2	2939	11	HSM801345	AL133074 Homo sapi	c 537	36	1.2	12597	39	HSCSE1G3	AF053645 Homo sapi
465	36	1.2	2966	5	A271171	A271171 Tyrosine ki	c 538	36	1.2	12661	11	AL133345	AL133345 Human DNA
466	36	1.2	2973	9	HUMSINEE	DI4548 Human DNA,	c 539	36	1.2	12728	12	RNU22062	U22062 Rattus norv
467	36	1.2	2990	50	AF138859	AF138859 Homo sapi	c 540	36	1.2	12786	12	AB01015S03	AB01015S03 Mus muscu
468	36	1.2	3015	10	MEDAZMRNA	X99971 Macaca fasc	c 541	36	1.2	13121	12	MMSRP20	X91636 M.musculus
469	36	1.2	3024	39	AF146568	AF146568 Homo sapi	c 542	36	1.2	13208	11	AB028893	AB028893 Homo sapi
470	36	1.2	3078	4	AF225906	AF225906 Xenopus l	c 543	36	1.2	13359	11	HSU89387	HSU89387 Human RNA p
471	36	1.2	3078	11	HSM801318	AL133047 Homo sapi	c 544	36	1.2	13728	11	AF005988	AF005988 Homo sapi
472	36	1.2	3113	12	AF056324	AF056324 Rattus no	c 545	36	1.2	14339	34	AE001414	AE001414 Plasmodiu
473	36	1.2	3133	11	HSM800556	AL080057 Homo sapi	c 546	36	1.2	14601	10	HSU84362	AL096762 Human DNA
474	36	1.2	3152	39	HSU78525	U78525 Homo sapien	c 547	36	1.2	14694	10	HUMDZAG	DI4034 Homo sapien
475	36	1.2	3169	11	HSM800953	AL117444 Homo sapi	c 548	36	1.2	15471	8	SPAC1B1	Z98532 S.pombe Chr
476	36	1.2	3171	4	AF115497	AF115497 xenopus l	c 549	36	1.2	15936	10	AB025285	AB025285 Homo sapi
477	36	1.2	3184	10	HS385E7	AL031720 Human DNA	c 550	36	1.2	16471	10	HS889J22A	AL049850 Human DNA
478	36	1.2	3209	34	AF181629	AF181629 Drosophil	c 551	36	1.2	17317	9	AP000250	AP000250 Homo sapi
479	36	1.2	3251	11	HSM800661	AL080148 Homo sapi	c 552	36	1.2	22738	9	AB014078	AB014078 Homo sapi
480	36	1.2	3256	50	AF208850	AF208850 Homo sapi	c 553	36	1.2	23996	71	AC027289	AC027289 Homo sapi
481	36	1.2	3262	33	DME011578	AJ011578 Drosophil	c 554	36	1.2	24159	50	AC005677	AC005677 Homo sapi
482	36	1.2	3336	8	AF118843	AF118843 Lycopersi	c 555	36	1.2	26106	39	AC006005	AC006005 Homo sapi
483	36	1.2	3357	9	HUMIRIT	M64936 Homo sapien	c 556	36	1.2	27170	9	AP000319	AP000319 Homo sapi
484	36	1.2	3363	33	CEMCE1A	X75564 C.elegans m	c 557	36	1.2	27755	9	AP000569	AP000569 Homo sapi
485	36	1.2	3441	11	AF057280	AF057280 Homo sapi	c 558	36	1.2	28247	10	HS362J20	AL049764 Human DNA
486	36	1.2	3451	5	E13998	AF13998 Cyanidium c	c 559	36	1.2	28835	10	HS444G9	Z98258 Human DNA s
487	36	1.2	3504	39	AF074606	AF074606 Homo sapi	c 560	36	1.2	29408	41	AC009804	AC009804 Homo sapi
488	36	1.2	3535	12	AF120492	AF120492 Rattus no	c 561	36	1.2	29848	39	U73023	U73023 Homo sapien
489	36	1.2	3551	10	HUM49DC62Z	L77073 Homo sapien	c 562	36	1.2	29952	10	HSV1164A6	Z69838 Human DNA s
490	36	1.2	3581	12	RNSTOP	X93495 R.norvegicu	c 563	36	1.2	30438	9	AP000299	AP000299 Homo sapi
491	36	1.2	3642	11	HSM801377	AL133021 Homo sapi	c 564	36	1.2	33595	32	HS312E8	AL032819 Homo sapi
492	36	1.2	3657	12	MMU51126	U51126 Mus musculu	c 565	36	1.2	33760	10	HSNFG9	Z69719 Human DNA s
493	36	1.2	3676	11	HSM802430	AL157448 Homo sapi	c 566	36	1.2	33905	55	AC010508	AC010508 Homo sapi
494	36	1.2	3700	2	HIU38617	U38617 Haemophilus	c 567	36	1.2	34352	11	U73638	U73638 Human Chrom
495	36	1.2	3738	11	HSM801291	AL133026 Homo sapi	c 568	36	1.2	35173	11	AL160413	AL160413 Human DNA
496	36	1.2	3765	11	HSM800941	AL117432 Homo sapi	c 569	36	1.2	35488	11	AC004589	AC004589 Homo sapi
497	36	1.2	3787	11	HSM801412	AL133565 Homo sapi	c 570	36	1.2	35730	10	HSN74G7	Z69715 Human DNA s
498	36	1.2	3798	9	HSSTREP	X58156 H.sapiens s	c 571	36	1.2	35843	10	HS108F12	Z49235 Human DNA s
499	36	1.2	3820	5	E02221	E02221 cDNA encodi	c 572	36	1.2	36003	51	AC020810	AC020810 Mus muscu
500	36	1.2	3851	5	E01614	E01614 DNA sequenc	c 573	36	1.2	36429	10	HS14B7	Z49258 Human DNA s
501	36	1.2	3851	5	E13364	E13364 Plasmid pyg	c 574	36	1.2	36565	11	HSU73648	Z49258 Human DNA s
502	36	1.2	3851	5	E13540	E13540 Seriola sp.	c 575	36	1.2	36649	11	HSU19D8	Z70689 Human DNA s
503	36	1.2	3970	12	AF027180	AF027180 Cavia por	c 576	36	1.2	36655	10	HS399E4	AL031721 Human DNA
504	36	1.2	4089	10	HUM25DC32	L78753 Homo sapien	c 577	36	1.2	36669	10	HS433G19	AL008733 Human DNA
505	36	1.2	4153	11	HSM801284	AL122121 Homo sapi	c 578	36	1.2	36781	9	AP000304	AP000304 Homo sapi
506	36	1.2	4178	10	HUMCPLA1	D38178 Human gene	c 579	36	1.2	37389	55	AC008974	AC008974 Homo sapi
507	36	1.2	4198	12	MUSHEX	J05277 Mouse hexok	c 580	36	1.2	37658	10	HS366L4	AL023494 Human DNA
508	36	1.2	4215	12	MMU73378	U73378 Mus musculu	c 581	36	1.2	37930	11	AC005564	AC005564 Homo sapi
509	36	1.2	4235	10	HSWEMD	Y10183 H.sapiens m	c 582	36	1.2	38194	39	AC004193	AC004193 Homo sapi
510	36	1.2	4295	9	AK000160	AK000160 Homo sapi	c 583	36	1.2	38357	10	HS462D8	AL022332 Human DNA
511	36	1.2	4417	11	AF098162	AF098162 Homo sapi	c 584	36	1.2	38437	9	AP000361	AP000361 Homo sapi
512	36	1.2	4493	11	HSM801283	AL122120 Homo sapi	c 585	36	1.2	38468	39	U51560	U51560 Homo sapien
513	36	1.2	4579	33	DME269539	AJ269539 Drosophil	c 586	36	1.2	38748	10	HSJ955M13	AL050404 Human DNA
514	36	1.2	4586	5	AR038747	AR038747 Sequence	c 587	36	1.2	38772	10	HSO19A	Z81310 Human DNA s
515	36	1.2	4586	5	AR059633	AR059633 Sequence	c 588	36	1.2	38968	11	AC005793	AC005793 Homo sapi
516	36	1.2	4740	10	HUMMAC1A	J03525 Human Mac-1	c 589	36	1.2	39000	10	HSN86D4	Z82250 Human DNA s
517	36	1.2	4827	11	AF059611	AF059611 Homo sapi	c 590	36	1.2	39019	10	D88270	D88270 Homo sapien
518	36	1.2	4978	5	I36871	I36871 Sequence 1	c 591	36	1.2	39021	39	AF052041	AF052041 Homo sapi
519	36	1.2	4978	23	E09332	E09332 cDNA encodi	c 592	36	1.2	39144	73	AC008980	AC008980 Homo sapi
520	36	1.2	4978	23	E10112	E10112 Cattle cDNA	c 593	36	1.2	39206	75	AC010649	AC010649 Homo sapi
521	36	1.2	5055	41	AC005843	AC005843 Homo sapi	c 594	36	1.2	39212	10	HSN21F1	Z94162 Human DNA s
522	36	1.2	5233	8	AF020288	AF020288 Arabidops	c 595	36	1.2	39238	39	AC002082	AC002082 Homo sapi
523	36	1.2	5301	11	AF061189	AF061189 Homo sapi	c 596	36	1.2	39239	11	AC005790	AC005790 Homo sapi
524	36	1.2	5307	11	AF040628	AF040628 Homo sapi	c 597	36	1.2	39265	55	AC010505	AC010505 Homo sapi
525	36	1.2	5321	11	HSM802199	AL137476 Homo sapi	c 598	36	1.2	39266	11	AC005791	AC005791 Homo sapi
526	36	1.2	5515	50	AF227198	AF227198 Homo sapi	c 599	36	1.2	39338	11	AC004644	AC004644 Homo sapi
527	36	1.2	5878	39	AF201468	AF201468 Homo sapi	c 600	36	1.2	39372	41	AC004176	AC004176 Homo sapi
528	36	1.2	8200	39	AF126531	AF126531 Homo sapi	c 601	36	1.2	39563	11	HSU73629	U73629 Human chrom
529	36	1.2	8212	12	MMHYLTK	X83972 M.musculus	c 602	36	1.2	39595	11	AC004509	AC004509 Homo sapi
530	36	1.2	8622	50	AF135026	AF135026 Homo sapi	c 603	36	1.2	39699	11	AC004170	AC004170 Homo sapi
531	36	1.2	8898	55	AC010551	AC010551 Homo sapi	c 604	36	1.2	39834	9	AF001410	AF001410 Homo sapi
532	36	1.2	9404	9	HUMBHSD	M38180 Human 3-bet	c 605	36	1.2	39838	39	AC007766	AC007766 Homo sapi
533	36	1.2	11597	9	AP000693	AP000693 Homo sapi	c 606	36	1.2	39842	9	AP000458	AP000458 Homo sapi
534	36	1.2	11807	9	D86990	D86990 Homo sapien	c 607	36	1.2	39976	11	AC005952	AC005952 Homo sapi
535	36	1.2	11967	10	HUMDNL1L	L40817 Homo sapien	c 608	36	1.2	40024	38	HSEWE5	Y08806 H.sapiens E

c 609	1.2	40160	31	AP001214	Homo sapi	682	1.2	57939	51	AC023608	Mus muscu
c 610	1.2	40205	11	AF045450	Homo sapi	683	1.2	58190	10	AC023608	Mus muscu
c 611	1.2	40209	67	AC026805	Homo sapi	684	1.2	59012	11	HS498124	U07562 Human ABL g
c 612	1.2	40307	9	AP001430	Homo sapi	685	1.2	59716	31	AP000637	U07562 Human ABL g
c 613	1.2	40607	11	AC005330	Homo sapi	686	1.2	59721	51	AC023413	U07562 Human ABL g
c 614	1.2	40616	55	AC016630	Homo sapi	687	1.2	59850	42	AC016452	U07562 Human ABL g
c 615	1.2	40649	11	AC003111	Human DNA	688	1.2	60129	39	AC006325	U07562 Human ABL g
c 616	1.2	40926	10	HSB10B1	Human DNA	689	1.2	60141	76	AC055858	U07562 Human ABL g
c 617	1.2	41133	11	AC004598	Homo sapi	690	1.2	60179	41	AF149786	U07562 Human ABL g
c 618	1.2	41191	55	AC009004	Homo sapi	691	1.2	60442	9	AP001342	U07562 Human ABL g
c 619	1.2	41495	11	AC004224	Homo sapi	692	1.2	60801	39	AC006561	U07562 Human ABL g
c 620	1.2	41617	11	AC004213	Homo sapi	693	1.2	61711	55	AC011530	U07562 Human ABL g
c 621	1.2	41617	11	AC004213	Homo sapi	694	1.2	62732	10	HS414D7	U07562 Human ABL g
c 622	1.2	41696	11	AC005932	Homo sapi	695	1.2	62732	10	HS414D7	U07562 Human ABL g
c 623	1.2	41702	11	AC004156	Homo sapi	696	1.2	62838	78	AC061995	U07562 Human ABL g
c 624	1.2	41726	70	AC027346	Homo sapi	697	1.2	62873	9	AP000238	U07562 Human ABL g
c 625	1.2	41907	39	AF050154	Homo sapi	698	1.2	62915	75	AC040984	U07562 Human ABL g
c 626	1.2	42179	39	HS076377	U07562 Human ABL g	699	1.2	63369	41	AC002413	U07562 Human ABL g
c 627	1.2	42345	55	AC010504	Homo sapi	700	1.2	64063	40	AL353578	U07562 Human ABL g
c 628	1.2	42547	39	AC023172	Homo sapi	701	1.2	64167	10	HSBA89B2	U07562 Human ABL g
c 629	1.2	42558	50	AC037199	Homo sapi	702	1.2	64167	10	HSBA89B2	U07562 Human ABL g
c 630	1.2	42601	11	AC004210	Homo sapi	703	1.2	64767	11	U69730	U07562 Human ABL g
c 631	1.2	42649	50	AC024077	Homo sapi	704	1.2	65346	67	AC026597	U07562 Human ABL g
c 632	1.2	42896	10	HS447C4	Human DNA	705	1.2	65730	42	AC014285	U07562 Human ABL g
c 633	1.2	42930	50	AF241732	Homo sapi	706	1.2	66109	9	HSBHCAPG	U07562 Human ABL g
c 634	1.2	43394	50	HSAC000359	Human cos	707	1.2	66739	69	AC008656	U07562 Human ABL g
c 635	1.2	43481	7	SCCHXIV43	Z46843 S cerevisia	708	1.2	66741	10	HS460J8	U07562 Human ABL g
c 636	1.2	43600	11	AC004209	Homo sapi	709	1.2	68663	73	AC036112	U07562 Human ABL g
c 637	1.2	43689	53	AC022516	Homo sapi	710	1.2	67053	9	AP001412	U07562 Human ABL g
c 638	1.2	43761	10	HS597B2	Human DNA	711	1.2	67298	31	AP000850	U07562 Human ABL g
c 639	1.2	43761	10	HS597B2	Human DNA	712	1.2	67331	55	AC010286	U07562 Human ABL g
c 640	1.2	43795	11	AC000022	Genomic s	713	1.2	67939	76	AC055843	U07562 Human ABL g
c 641	1.2	43978	74	AC010617	Homo sapi	714	1.2	67994	39	AC006981	U07562 Human ABL g
c 642	1.2	44210	55	AC016629	Homo sapi	715	1.2	68245	42	AC018626	U07562 Human ABL g
c 643	1.2	44348	11	AC006046	Homo sapi	716	1.2	68581	9	AP000563	U07562 Human ABL g
c 644	1.2	44496	11	AC004760	Homo sapi	717	1.2	68730	39	HSU95739	U07562 Human ABL g
c 645	1.2	44500	73	AC011522	Homo sapi	718	1.2	69015	39	HSU95739	U07562 Human ABL g
c 646	1.2	44837	55	AC011544	Homo sapi	719	1.2	69017	10	HS410I8	U07562 Human ABL g
c 647	1.2	44890	50	AC005778	Homo sapi	720	1.2	69142	7	AB012239	U07562 Human ABL g
c 648	1.2	44890	50	AC005778	Homo sapi	721	1.2	69279	76	AC046187	U07562 Human ABL g
c 649	1.2	44978	50	AF241731	Homo sapi	722	1.2	69517	73	AC036117	U07562 Human ABL g
c 650	1.2	45084	11	AC003005	Human DNA	723	1.2	69529	55	AC012632	U07562 Human ABL g
c 651	1.2	45337	10	HS10618A	Human DNA	724	1.2	69572	11	AC004769	U07562 Human ABL g
c 652	1.2	45433	41	AC002095	Homo sapi	725	1.2	69684	39	AC006475	U07562 Human ABL g
c 653	1.2	45550	39	AC006292	Homo sapi	726	1.2	70165	74	AC037463	U07562 Human ABL g
c 654	1.2	46201	39	AC006139	Homo sapi	727	1.2	70488	42	AC015519	U07562 Human ABL g
c 655	1.2	46275	11	AC003107	Human DNA	728	1.2	70739	67	AC026366	U07562 Human ABL g
c 656	1.2	46534	10	HS1057D18	Human DNA	729	1.2	70851	39	AP000269	U07562 Human ABL g
c 657	1.2	48092	76	AC053543	Homo sapi	730	1.2	70932	9	AP000269	U07562 Human ABL g
c 658	1.2	48425	51	AC023310	Homo sapi	731	1.2	71534	76	AC046188	U07562 Human ABL g
c 659	1.2	48545	71	AC011515	Homo sapi	732	1.2	71613	73	AC036139	U07562 Human ABL g
c 660	1.2	48566	69	AC025172	Homo sapi	733	1.2	72289	11	HSU91115	U07562 Human ABL g
c 661	1.2	48645	11	AC005575	Homo sapi	734	1.2	72852	31	AP000686	U07562 Human ABL g
c 662	1.2	49261	11	HSBA209A2	Human DNA	735	1.2	74155	11	AC004656	U07562 Human ABL g
c 663	1.2	49499	9	AF184110	Homo sapi	736	1.2	74195	72	AC034125	U07562 Human ABL g
c 664	1.2	50188	9	AP000047	Homo sapi	737	1.2	74549	10	HSJ890015	U07562 Human ABL g
c 665	1.2	50508	39	AC004940	Homo sapi	738	1.2	74729	10	HSJ890015	U07562 Human ABL g
c 666	1.2	50592	31	AP001236	Homo sapi	739	1.2	73231	39	AC005078	U07562 Human ABL g
c 667	1.2	50604	39	AC005074	Homo sapi	740	1.2	73842	57	AC025657	U07562 Human ABL g
c 668	1.2	50768	39	AF095901	Homo sapi	741	1.2	74116	42	AC016110	U07562 Human ABL g
c 669	1.2	50768	39	AF095901	Homo sapi	742	1.2	74155	11	AC004656	U07562 Human ABL g
c 670	1.2	51859	10	HSJ1193N1	Human DNA	743	1.2	74195	72	AC034125	U07562 Human ABL g
c 671	1.2	53168	10	HS109G6	Human DNA	744	1.2	74549	10	HSJ890015	U07562 Human ABL g
c 672	1.2	53897	11	AC002037	Human Chr	745	1.2	74729	10	HSJ890015	U07562 Human ABL g
c 673	1.2	55520	10	HSU4488M4	Human DNA	746	1.2	75270	11	AF023268	U07562 Human ABL g
c 674	1.2	55917	39	AC006002	Human DNA	747	1.2	75270	11	AF023268	U07562 Human ABL g
c 675	1.2	56730	39	AC006314	Homo sapi	748	1.2	75976	73	AC026633	U07562 Human ABL g
c 676	1.2	56827	39	HSBHC3A5	Homo sapi	749	1.2	76079	39	AC004942	U07562 Human ABL g
c 677	1.2	56846	46	AC010850	Homo sapi	750	1.2	76815	75	AC044789	U07562 Human ABL g
c 678	1.2	56913	10	HS67C13	Human DNA	751	1.2	76822	72	AC032041	U07562 Human ABL g
c 679	1.2	56997	32	CNS01RG2	Human DNA	752	1.2	76986	39	AC006578	U07562 Human ABL g
c 680	1.2	57166	39	AC009320	Homo sapi	753	1.2	76990	31	AP001770	U07562 Human ABL g
c 681	1.2	57166	39	AC009320	Homo sapi	754	1.2	77183	10	HS299C21	U07562 Human ABL g

755	36	1.2	77879	70	AC027403	AC027403 Homo sapi	828	36	1.2	90357	61	AC023880	AC023880 Homo sapi
756	36	1.2	78693	10	HSJD475B7	AL050306 Human DNA	829	36	1.2	90433	40	AL139289	AL139289 Homo sapi
757	36	1.2	79085	39	AC006221	AC006221 Homo sapi	830	36	1.2	90618	57	AC018667	AC018667 Homo sapi
758	36	1.2	79205	75	AC008598	AC008598 Homo sapi	c 831	36	1.2	90744	10	HSJ646B12	AL096776 Human DNA
759	36	1.2	79716	11	AC004502	AC004502 Homo sapi	c 832	36	1.2	90815	67	AC002559	AC002559 Homo sapi
760	36	1.2	79868	70	AC027408	AC027408 Homo sapi	c 833	36	1.2	90879	67	AC026356	AC026356 Homo sapi
761	36	1.2	80141	9	HSEWSGAR	Y07848 Homo sapien	c 834	36	1.2	91047	67	AC026430	AC026430 Homo sapi
762	36	1.2	80165	71	AC027745	AC027745 Homo sapi	c 835	36	1.2	91128	31	AP001826	AP001826 Homo sapi
763	36	1.2	80168	39	AC002078	AC002078 Human BAC	c 836	36	1.2	91365	10	HS116F5	Z93244 Human DNA s
764	36	1.2	80209	32	CNS01DVI	AL136294 Homo sapi	c 837	36	1.2	91368	39	AC005994	AC005994 Homo sapi
765	36	1.2	80331	39	AC004994	AC004994 Homo sapi	c 838	36	1.2	91511	11	AL133404	AL133404 Human DNA
766	36	1.2	80512	52	AC007834_3	Continuatiion (4 of	c 839	36	1.2	91634	39	AC005008	AC005008 Homo sapi
767	36	1.2	80518	31	PFMAL13PA	AL109815 Plasmodiu	c 840	36	1.2	91638	41	AC002489	AC002489 Mus muscu
768	36	1.2	80858	39	AC007360	AC007360 Homo sapi	c 841	36	1.2	91638	41	AC002489	AC002489 Mus muscu
769	36	1.2	80869	10	HSJD454N4	AL096794 Human DNA	c 842	36	1.2	91767	39	AC007748	AC007748 Homo sapi
770	36	1.2	81290	51	AC022597	AC022597 Homo sapi	c 843	36	1.2	91886	32	AL133472	AL133472 Homo sapi
771	36	1.2	81340	43	AC021612	AC021612 Homo sapi	c 844	36	1.2	92073	72	AC008968	AC008968 Homo sapi
772	36	1.2	81374	11	HSJ300H18	AL132799 Human DNA	c 845	36	1.2	92357	10	HSJ106901	AL049633 Human DNA
773	36	1.2	81404	52	AC008180_3	Continuatiion (4 of	c 846	36	1.2	92636	39	AC005251	AC005251 Homo sapi
774	36	1.2	81479	70	AC026704	AC026704 Homo sapi	c 847	36	1.2	92651	39	AC002527	AC002527 Human BAC
775	36	1.2	81521	39	AC006255	AC006255 Homo sapi	c 848	36	1.2	92797	11	AF038458	AF038458 Homo sapi
776	36	1.2	81561	9	AB023057	AB023057 Homo sapi	c 849	36	1.2	92817	11	HSJ393D12	AL132776 Human DNA
777	36	1.2	81807	72	AC011433	AC011433 Homo sapi	c 850	36	1.2	92969	11	HSJ687F11	AL117334 Human DNA
778	36	1.2	82116	9	AP000338	AP000338 Homo sapi	c 851	36	1.2	93273	39	AC004084	AC004084 Homo sapi
779	36	1.2	82237	50	AC003677	AC003677 Human P1	c 852	36	1.2	93273	32	AL135912	AL135912 Homo sapi
780	36	1.2	82426	76	AC055885	AC055885 Homo sapi	c 853	36	1.2	94106	11	AC005667	AC005667 Homo sapi
781	36	1.2	82690	31	AP001128	AP001128 Homo sapi	c 854	36	1.2	94148	41	HSJ858B6	AL118511 Homo sapi
782	36	1.2	83023	73	AC036231	AC036231 Homo sapi	c 855	36	1.2	94308	39	AC005480	AC005480 Homo sapi
783	36	1.2	83079	39	AC004538	AC004538 Homo sapi	c 856	36	1.2	94320	39	AC002564	AC002564 Human BAC
784	36	1.2	83514	31	AP001129	AP001129 Homo sapi	c 857	36	1.2	94320	39	AC002564	AC002564 Human BAC
785	36	1.2	83534	69	AC008933	AC008933 Homo sapi	c 858	36	1.2	94336	39	AC005250	AC005250 Homo sapi
786	36	1.2	84126	10	HS125N5	AL008728 Human DNA	c 859	36	1.2	94374	72	AC008384	AC008384 Homo sapi
787	36	1.2	84307	77	AC055742	AC055742 Homo sapi	c 860	36	1.2	94715	32	HSJ575L21	AL096841 Homo sapi
788	36	1.2	84412	10	HS973M2	AL033533 Human DNA	c 861	36	1.2	94960	39	AC000066	AC000066 Homo sapi
789	36	1.2	84464	10	HS782D21	AL023399 Human DNA	c 862	36	1.2	95038	72	AC011430	AC011430 Homo sapi
790	36	1.2	84633	39	AC005661	AC005661 citb_54_O	c 863	36	1.2	95267	10	HSJ1J24	AL024474 Human DNA
791	36	1.2	84871	10	HSJD5AG6	AL121588 Human DNA	c 864	36	1.2	95367	39	AC004001	AC004001 Human BAC
792	36	1.2	84912	11	AF024533	AF024533 Homo sapi	c 865	36	1.2	95433	67	AC026437	AC026437 Homo sapi
793	36	1.2	85058	39	AC004989	AC004989 Homo sapi	c 866	36	1.2	95855	11	HSAC000115	AC000115 Human BAC
794	36	1.2	85084	72	AC011426	AC011426 Homo sapi	c 867	36	1.2	95956	39	AC006336	AC006336 Homo sapi
795	36	1.2	85135	11	HS136017	Z72001 Human DNA s	c 868	36	1.2	96256	10	HSJ177I5	AL022315 Human DNA
796	36	1.2	85579	39	AC004744	AC004744 Homo sapi	c 869	36	1.2	96276	10	HSJ1175I6	AL049538 Human DNA
797	36	1.2	85654	10	HSP373C6	AL022393 Homo sapi	c 870	36	1.2	96308	39	AF190465	AF190465 Homo sapi
798	36	1.2	85863	72	AC010464	AC010464 Homo sapi	c 871	36	1.2	97060	55	AC010350	AC010350 Homo sapi
799	36	1.2	85931	12	AC007049	AC007049 Mus muscu	c 872	36	1.2	97392	10	HSBA150A6	AL096770 Human DNA
800	36	1.2	86049	80	AC006585	AC006585 Homo sapi	c 873	36	1.2	98117	11	AC003688	AC003688 Homo sapi
801	36	1.2	86394	32	HSJ84J3	AL117343 Homo sapi	c 874	36	1.2	98117	11	AC003688	AC003688 Homo sapi
802	36	1.2	86446	55	AC008754	AC008754 Homo sapi	c 875	36	1.2	98183	67	AC026448	AC026448 Homo sapi
803	36	1.2	86516	50	AC009493	AC009493 Homo sapi	c 876	36	1.2	98447	10	HSJ23H9	AL078622 Human DNA
804	36	1.2	86684	50	AC010605	AC010605 Homo sapi	c 877	36	1.2	98562	50	AC008969	AC008969 Homo sapi
805	36	1.2	86828	43	AC020569	AC020569 Homo sapi	c 878	36	1.2	98573	39	AF095725	AF095725 Homo sapi
806	36	1.2	86896	10	HS93N13	Z84489 Human DNA s	c 879	36	1.2	98684	32	AL133516	AL133516 Homo sapi
807	36	1.2	86969	79	AC063376	AC063376 Homo sapi	c 880	36	1.2	98697	39	AC004854	AC004854 Homo sapi
808	36	1.2	87013	10	HS528L19	Z84486 Human DNA s	c 881	36	1.2	98732	41	HSJ46801	AL109923 Homo sapi
809	36	1.2	87114	11	AC005353	AC005353 Homo sapi	c 882	36	1.2	98758	10	HSJ23H9	AL078622 Human DNA
810	36	1.2	87733	72	AC015600	AC015600 Homo sapi	c 883	36	1.2	98828	10	HSJ388E23	AL049552 Human DNA
811	36	1.2	88013	73	AC008402	AC008402 Homo sapi	c 884	36	1.2	99370	39	AC005057	AC005057 Homo sapi
812	36	1.2	88050	9	AP001421	AP001421 Homo sapi	c 885	36	1.2	99667	32	AL157376	AL157376 Homo sapi
813	36	1.2	88071	11	AC002316	AC002316 Homo sapi	c 886	36	1.2	99674	11	CNS01DSP	AL121838 Human chr
814	36	1.2	88133	74	AC026771	AC026771 Homo sapi	c 887	36	1.2	99902	10	HSJ59B16	AL032822 Human DNA
815	36	1.2	88215	39	AC006963	AC006963 Homo sapi	c 888	36	1.2	99935	32	AL136975	AL136975 Homo sapi
816	36	1.2	88217	41	AC006407	AC006407 Homo sapi	c 889	36	1.2	100000	9	AB020867	AB020867 Homo sapi
817	36	1.2	88882	10	HS448E20	Z97196 Human DNA s	c 890	36	1.2	100000	9	AP000011	AP000011 Homo sapi
818	36	1.2	89045	73	AC009185	AC009185 Homo sapi	c 891	36	1.2	100000	9	AP000013	AP000013 Homo sapi
819	36	1.2	89218	39	AC005902	AC005902 Homo sapi	c 892	36	1.2	100000	9	AP000018	AP000018 Homo sapi
820	36	1.2	89240	67	AC026420	AC026420 Homo sapi	c 893	36	1.2	100000	9	AP000020	AP000020 Homo sapi
821	36	1.2	89292	9	AP001252	AP001252 Homo sapi	c 894	36	1.2	100000	9	AP000030	AP000030 Homo sapi
822	36	1.2	89323	9	AP000221	AP000221 Homo sapi	c 895	36	1.2	100000	9	AP000033	AP000033 Homo sapi
823	36	1.2	89328	10	HS398C22	Z93784 Human DNA s	c 896	36	1.2	100000	9	AP000045	AP000045 Homo sapi
824	36	1.2	89800	69	AC010474	AC010474 Homo sapi	c 897	36	1.2	100000	9	AP000052	AP000052 Homo sapi
825	36	1.2	90082	11	AF045555	AF045555 Homo sapi	c 898	36	1.2	100000	9	AP000071	AP000071 Homo sapi
826	36	1.2	90220	10	HSJ734P14	AL049650 Human DNA	c 899	36	1.2	100000	9	AP000077	AP000077 Homo sapi
827	36	1.2	90299	11	HS72E17	AL033523 Human DNA	c 900	36	1.2	100000	9	AP000080	AP000080 Homo sapi

NPYRTEIKEELEEDRIRTEQTLFDLYKYNSSYTRQAGARRSSRDLGAFPHPLQRL
RTPPPYSGRTARSSSSVRRDNNPQVDRKDWKIGFOLCQNONKSDFCYQSYSSGVDAVR
EYRFRHYINILSRLSDTSPALREELGNFIETCRFNOAPCNOANTSKFHPHYGNCYT
FNDKNSLNWSPGCVNNGLSLTLRTEONDFIPLLSVTGARVMVHGODEFMDDG
GFNLPGVETSISMKEALDSLGGNYGDCTEGSDVPVKNLYPSKYTOQVCILHSCFE
NMKKCGCAYIYFYPKGVDFCDYRKQSSWGCYYKLOGAFSLDSLGFCSKRKPCSV
INYLUSAGYSRMPSSQDWIFEMISLQNNYITINKRNGVAKLNIFFKELNTKNS
PSVWVLSNLGSSQSLWFGSVLSVEMADVIEDLLVITLLMLLRFRSRYWSPGR
GARGAREVASTPFPSPFRPHPTSPPLPQQGTMTPPLALTAPPAYATLGPSPAPP
LDSAAPDCSACALAL"

BASE COUNT 684 a 983 c 753 g 697 t

ORIGIN

Query Match		100.0%;	Score 3117;	DB 12;	Length 3117;		
Best Local Similarity		100.0%;	Pred. No. 0;				
Matches 3117;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	cegggtgcacccacgctgcgagggcagcctgggagtgogggcagcgtcccgacagccc	60				
Db	1	CCGGGTGAGCCACCGGTCGAGGGCAGCCTGGGATGCGGGCAGCGTCCCGGACAGCCC	60				
QY	61	cattcgcttcacgtaataatgatgctggaccacacagagccctgagctcaacattgac	120				
Db	61	CATTCTGCGCTTACGGTAATGATGCTGGACACACAGAGCCCTGAGCTCAACATTTGAC	120				
QY	121	ctagaccttcacgcctcccaactgcctaaaggggtccatgaagggaaccaaatcaagag	180				
Db	121	CTAGACCTTCACGCCCTCCAACTCGCCTAAGGGGTCCATGAAGGGCAACCAATTCAGGAG	180				
QY	181	caagacccctgtctcctcagcccatcgaaggactgggaaggggggaacaacgtgaagag	240				
Db	181	CAAGACCCTTGTCTCTCAGCCCATGCAAGGACTGGGAAGGGGACAAACGTGAAGAG	240				
QY	241	caaggcctgggcccgaacctcagcaccccggcagcccaacagagaggaggaggaactg	300				
Db	241	CAGGGCTTGGCCCGGAACCCCTCAGCACCCCGGACGCCACCGAGGAGGAGGACGACTG	300				
QY	301	atgaattccacgcctcctacgggagctctccagttcttctgcaacaacacaccatc	360				
Db	301	ATTGAATTCACGCCTCTTACCGGGAGCTCTTCCAGTTCTTCTGCAACAACACCCATC	360				
QY	361	cacggggccatccgcctgggtgctccaaacacaaacccgeatgaagcggcctctcggcg	420				
Db	361	CACGGGCCATCCGCCTGGTGTCTCCAAACACAAACCGCATGAAGCGGCTTCTGGGCG	420				
QY	421	gtcgtgtgctgtgacacttcggcatgatgtactggcagcttcgcttgcgttcgagag	480				
Db	421	GTGCTGTGGCTGTGCACCTTTCGGCATGATGTACTGGCAGTTTCGCTTGTTCGAGGAG	480				
QY	481	tactcagctaccagctgagcctcaactcaactcaactcagacagctggtctccct	540				
Db	481	TACTCTAGCTTACCAGTGAGCCTCAACATCAACCTCAATTCAGACAACAGCTGTCTCCCT	540				
QY	541	gccgtcactgtctgcacccttaactctacagatacactgaactgaataaagaggagctggaa	600				
Db	541	GCCGTCACTGTCTGCACCCCTTAATCTTACAGATACACTGAATTAAGAGAGGAGCTGGAA	600				
QY	601	gagctgacgcatcagcagcagcagcgtttttgactgtacaaaatacaactcttctac	660				
Db	601	GAGCTGAGCGCATCAGGAGCAGAGCGCTTTTGACTTGTACAAATACAACTCTTCTCTAC	660				
QY	661	actcgcagcgtgggcccagcgcgcagctcccgagcctcctcgtggtgttttcccgac	720				
Db	661	ACTCGCAGGCTGGGGCCCGACCCGAGCTTCCCGGACCTCCTGGGTGCTTTCCCGCAC	720				
QY	721	ccctcagcgcctgcgactccactccgctactccgcccgcgcgcgcgcgcgcgcgcgcgc	780				
Db	721	CCCTCGACGGCTTGGCGACTCCACCTCCGCCCTACTCCGCCCGCACGGGCCGACGGGG	780				
QY	781	tcttcagcgtacgcgacacaaatcccccaagtggagccggaagactcggaactcggttc	840				
Db	781	TCTTCCAGCGTACGCGACACAATCCCCAAGTGGACCGGAAGGACTGGAAGATCGGCTTC	840				

QY	841	caactgtgcaaccagaaacaatcagactgtttctaccagacatactcctctggggtggat	900
Db	841	CAACTGTGCAACAGCAACAATCAGACTGTTCTTACCAGACATACTCTCTGGGGTGAT	900
QY	901	gcagtgaggagtggtaccgcttccatcatcaacaatctgtccagacgtcggaacac	960
Db	901	GCAGTGAGGAGTGGTACCGCTTCCATTATACATCAACATTTCTGTCCAGACTGTCGAGACC	960
QY	961	tcgccgcctctagagaagaagccctgggcaacttcattcctcactgctcactcaacag	1020
Db	961	TCGCCGCTCTAGAGAAGAAGCCCTGGCAACTTCATCTTCACTGTGCGCTTCAACAG	1020
QY	1021	gccccctgcaaccaggcggaattattccaagttccaccaccccatgtacgggaactgctac	1080
Db	1021	GCCCCCTGCAACAGCGGAATTATTCCAAGTCCACCACCCCATGTACGGGAACGTGCTAC	1080
QY	1081	acttcaatgacaagaacaactccatctctggatgtctctccatgctcgtgagtaacaat	1140
Db	1081	ACTTCAATGACAAGAACAACCTCCAATCTCTGGATGTCTCCATGCTGGAGTCAACAAT	1140
QY	1141	ggtttgctccctgacactgcacagacagagaactccatccctcgtcgtccacagt	1200
Db	1141	GGTTGTGCTTGCCTGACACTGCGCACAGACGAATGACTTCATFCCCCCTGCTGTCCACAGTG	1200
QY	1201	acgggggcccagggtgagtggtcaggtcaggtcaggtcgtcctttatggatgagtggc	1260
Db	1201	ACGGGGGCCAGGGTGAAGTGGTGCATGGTGCAGTGAAGGCTGCTTATGGATGATGGTGGC	1260
QY	1261	tcaacttgaggcctggcgtggagacctccatcagtagagaaggagccctgggacgc	1320
Db	1261	TCAACTTGAGGCTGGCGTGGAGACCTCCATCATAGTAGAAAAGGAAGCCCTGGACAGC	1320
QY	1321	ctcggggaataattacggcgactgactgagaatggttagcgatgcccggtcaagaacct	1380
Db	1321	CTCGGAGGAATTACGGCGACTGTACTGAGAATGGTAGCGATGTCCCGGTCAAGAAGCTT	1380
QY	1381	taccttccagtatcacacagcaggtgtgattcaactcctgctccagggagaaacatgac	1440
Db	1381	TACCTTCCAAGTATACACAGCAGGTGTGCATTCACCTCTCTTCCAGGAGAACATGATC	1440
QY	1441	aagaagtgtgctgctacatctctaccctaaagcccaaggagggttaggtctctgac	1500
Db	1441	AAGAAGTGTGCTGTGCTACATCTTACCTTAAGCCCAAGGAGTGTAGTTCTGTGAC	1500
QY	1501	taccgaagcagagcctcctgggctattgattataaaactgcaaggcgcctctccttg	1560
Db	1501	TACCGAAGCAGAGCTCCTGGGCTATTGCTATTATTAACATGCAGGCGCTTCTCCTTG	1560
QY	1561	gacagcctgggctgtttctcgaagtgtcgaagccttgtagtgatcaactacaactc	1620
Db	1561	GACAGCCTGGGCTGTTTCTCCAAGTGTGGAAGCCTTGTAGTGTGATCACTACAACTC	1620
QY	1621	cttgccggtcactcagcgtggccatctgtgaagtcaccagattggattcttcgagatcgt	1680
Db	1621	CTGCGGCTACTCACGGTGGCCATCTGTGAAGTCCCGAGGATTGGATCTTTCGAGATGCTG	1680
QY	1681	tcottgcagaaacattacactattacaacaaagaacagggtgcaagcgtcaacac	1740
Db	1681	TCCTTGAGAACAAATTACACTATTAAACAACAAAGAGAGTGTGCAAGCTCAACATC	1740
QY	1741	tcttcaagagcgtgaactataaaactaatcggagctcctctcgtcagatggtcagc	1800
Db	1741	TTCTTCAAGGAGCTGAACATAAAACATAATTCGGAGTCTCCTTCTGTCAGATGGTCAAG	1800
QY	1801	ctcctgtccaaacctgggcagccagtgagcctgtgggttggtcgtcgtgctctcgtg	1860
Db	1801	CTCCTGTCCAACTGGGCAGCCAGTGGAGCCTGTGGTTTGGCTGCTCGGTCTCTGTG	1860
QY	1861	gtggagatggcgacgtcatctctcgacctcctggtcatcacactctcatgctgctacgc	1920
Db	1861	GTGGAGATGGCGAGCGTCATCTTCGACCTCCTTGGTCATCACACTTCTCATCTGCTCTACGC	1920

```
QY 1921 cggttccggagcggtactgtctccaggacgagggccagggtgcccaggaggtggcc 1980
DB 1921 CGGTTCCGGAGCGGTACTGCTCTCCAGGAGAGGGGCCAGGGGTGCCAGGAGGTGCC 1980
QY 1981 tccactccagcttctctctccctcccgctcccgcttctctctccactccatccatccaccacct 2040
DB 1981 TCCACTCCAGCTTCTCTCTCCGCTCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2040
QY 2041 tctttgcccagcagggcgtacgacctccccctgcccctgcccctgacagccctccactgacctat 2100
DB 2041 TCTTTGCCCCAGCAGGCGATGACCCCTTCCCTTGCCCTGACAGCCCTTCCACCTGCCCTAT 2100
QY 2101 gctacttagccccagtgccccctccactgacctctcgccgctgactgtctctgacctgt 2160
DB 2101 GCTACTCTAGCCCCAGTGCCCCCTTCCACTGGAGCTCTCGCGGCCCTGACTGTCTGCCCCTGT 2160
QY 2161 gccctggcgcgtctctgagagagagagagagagagagagagagagagagagagagagagagag 2220
DB 2161 GCCCTGGCGCGCTCTGAGAGAGAGAGAGAGATCTCTCACCCAGCCCTGAGCTCCCTTGT 2220
QY 2221 aaactgatgatctcacttcacgtcacatctccccagcgccccgcgcgagccctctctgtg 2280
DB 2221 AAACGTGATGATATCTCATTACACATCTTCCCCAGCGGCCCGCCGAGCCCTCTCTGTG 2280
QY 2281 gtgtgctgctgagggagacaggtgagtaaggggggccaggaagcgtccagagagacagg 2340
DB 2281 GTGTGCTGCTGAGGAGACAGGCTGAGTAAGGGGCCAGGAAGCCGTCCAGAGGACAGG 2340
QY 2341 ggctaagatgtctcagagctgacctgccccctgctctctgaaacactgacctccacaagca 2400
DB 2341 GGCTAAGGATCTGCTAGAGCTGCCCTGCCCTGCTTCTGAACACTGCTTCCACAAGCA 2400
QY 2401 cagacaagtctctctccctcgatcagcccaagcagacttgagctctgacaaggaac 2460
DB 2401 CAGACAAGTTCCCTCTTCCCTCGGATCAGCCAGCCAGACTGGAGCTCTGACAAAGAAC 2460
QY 2461 ttctctggagagacaaagaaataagaaagacatgcactcctacccccgtgacc 2520
DB 2461 TTTCTCTGGAGAGACCAAGGAACAATAAGAACAGCCTACGACTCTCTACCCCGTGACC 2520
QY 2521 agaggtggcctggcctcactgctctcaggagacactgatgtctgctcctctccttgaact 2580
DB 2521 AGAGGCTGGCGTGGCTCAGTCCCTTCAAGGACACTGATGTCTGCTCTCTCTCTTGAAT 2580
QY 2581 tgggtggggaaccccccaaaagccccctgttagttctttggcaattcccccccctc 2640
DB 2581 TGGTGGGGAACCCCAAAAGCCCCCTTGTGTAGTTCTTTGGCAATCCCCCTCCCTC 2640
QY 2641 actcctcaggtgggactagtagtaagacagacagaggtaggcctggtgtctctctcat 2700
DB 2641 ACTCCTCAGGTTGGGACTAGAGTAAGACAGACCAGGGTAGGCCCTGGTGTCTCTCTCAT 2700
QY 2701 tctcaagactctcattgtatcacaccataccccagctgcctctgtggtgtgtgtctcttc 2760
DB 2701 TCTCAAGACTCTCTATTGTATACACCATACCCAGTGCCTCTGTGGTGTGTGTCTCTCTC 2760
QY 2761 tttctgttctcagcctggaagtcttttcaactgtccctcagaacacacagctatgca 2820
DB 2761 TTTCTGTCTTCTCAGCTTGAAGTTCTTTTACATGTCCCTCAGAACACACAGTCTATGCA 2820
QY 2821 tcccttagaaacctgcccagagcaccatttctttgtgaaactcagttaccctgttatgt 2880
DB 2821 TCCCTTAGAAACCTTGCCAAAGGACCATTTCTTTTGTGAACCTCAGTTACCTGTTATGT 2880
QY 2881 ctccccagaattgtgocctctccctccacccccacagtgactgtattattatgttcaac 2940
DB 2881 CTCCCCAGAATTGTGCGCCCTCTCCCTTCACCCCCACAGTGTACTGTATTATATTGTTTCA 2940
QY 2941 attcttttgtctgctccctgggtaaacactgagctcctctgtgtggtcaggatgagatt 3000
DB 2941 ATTCTTTGTGCTGCTCCCTGGGTAAACTGAGCTCCCTGTGTGGTCAGGGATGAGATT 3000
QY 3001 tgcctgttttgtatccttcccgctcctagcccgactctcccaacttgggacgggttaggcaggt 3060
```

```
DB 3001 TGCCTGTTTGTATCTTCCGGTCTAGCCAGTCTCCACTTGGGACGGGTAGSCAGGT 3060
QY 3061 actcaataaaggctgttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 3061 ACTCAATAAAGGCTTGTTCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3117

RESULT 2
RNASNAC 3081 bp mRNA ROD 07-DEC-1993
LOCUS R. norvegicus mRNA for amiloride sensitive Na+ channel protein.
DEFINITION X70521
ACCESSION X70521
VERSION X70521.1 GI:433909
KEYWORDS amiloride sensitive; Na+ channel protein.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 3081)
AUTHORS Lingueglia, E., Volley, N., Waldmann, R., Lazdunski, M. and Barbry, P.
TITLE Expression cloning of an epithelial amiloride-sensitive Na+
channel. A new channel type with homologies to Caenorhabditis
elegans degenerins
JOURNAL FEBS Lett. 318 (1), 95-99 (1993)
MEDLINE 93170495
REFERENCE 2 (bases 1 to 3081)
AUTHORS Barbry, P.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-1993) P. Barbry, IPMC UPR411 CNRS, 660 route des
Lucioles 06560, Sophia Antipolis, Cedex, FRANCE
FEATURES
source
1..3081
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="distal colon"
/cell_type="epithelial"
54..2153
/codon_start=1
/product="amiloride sensitive Na+ channel protein"
/protein_id="CAA49916.1"
/db_xref="GI:433910"
/db_xref="SWISS-PROT:P37089"
/translaton="MMLDHTRAPELINIDLHASNPKGSMKGNQFKEQDPCPQPMQ
GLGKDKREOGIGPEPSAPROPTEEELIEFHSYRELFQFFCNNTTHGAILRLVC
SKNRMKTAFWAVLWLCITFGMYWQFALLFEEVLSYPSVLNINLSNKLVPAYTVCT
LNPVRYTEIKEELELDRIETQFLDYKYNSSYTRQAGARRRSRLDLGAFPHPLQOR
LRTPPPYSGRTARSGSSSVRDNNPOVDKDWKIGFOLCQNKSDCFQVTVSSGVDAY
REWYRFHYINILSLSDTSPALEEALGNFTFCRFNQAPCNQANYSKFHHPMYGNCY
TFNDKNSNLWSSMPCVNNGLSLTLTEQNDIFPLLSTVTGARVYHVGDEPAPMDQ
GFFNLRPGETSISMRKEALDSLGNGDCTENGSDYFVKNLXPSKYTQQVCITHCQFQ
ENMIKGCQAYIFYPKPKGVFCFYRKQSSWGYYCYLQGLAFSLDSLGCSFKRKPQS
VINYLKSGYSRWPVKVSDQWIFEMLSLNNYTINNRNGVAKLINERFELNLYTNSF
SPSVTVSLILNLGWSMLWFGSSVLSVEMAEELIFDLLVITLLMLLRPRSRVYNSPG
RGARGAREVASTPASSPFRCPHTSPFPLSQQGMTPPLALTAPPAYATLQPSAP
PLDGAAPDCSACALAL"
BASE COUNT 669 a 974 c 742 g 696 t
ORIGIN

Query Match 89.7%; Score 2796; DB 12; Length 3081;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3076; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 27 cagctcggtatcgggcacggtcccgagacgccccattctgccttcacgcqctaagtatgct 86
DB 2 CAGCCTGGGATCGGGCACGGTCCCGGACAGCCCCATTCTGCCTCACGCTAATGATGCT 61
QY 87 ggacacaccagagccctcgactcaacattgaactagacctcaacgctcccaactcgcc 146
DB 62 GGACACACAGAGCCCTCGAGTCAACATTGACCTAGACCTTACGGCTTACGGCTCCACACTCGCC 121
```


QY 467 tgctgttcgagagtagtacctcagctaccagtgtagcctcagctcaacatcaacctcaattcagaca 526
DB 421 TGCTGTTCGAGGAGTACCTCAGCTACCCAGTGAGCCTCAACATCAACCTCAATTCCAGACA 480
QY 527 agctggcttcctccgcgtcaactgtctgcacccttaactccttacagatacaactgaatta 586
DB 481 AGCTGGCTCTTCCTCGCGTCACTGTCTGCACCTTTAATCCTTTACAGATACACTGAATAA 540
QY 587 aagagagctggaagagctggaccgcgcacacgagcagacgctcttttgacttgatacaaat 646
DB 541 AAGAGGAGCTGGAAGAGCTGGACCGCATCAGGAGCAGACGCTTTTGTGACTTGTACAAAT 600
QY 647 acaactcttcctacactcgcagggctggggccgcagccgcgcagctcccgagacctcctgg 706
DB 601 ACAACTCTTCTTACACTGCCAGGCTGGGGCCGCGCAGCTCCCGCGACCTCCTCG 660
QY 707 gtgtttccgcgaacccccctgcagcgctgcgaactcccaactcgcactactcgcggcgca 766
DB 661 GTGCTTTCCCGCACCCCTCTGCAGCGCTCGCACCTCCACCTCCGCCCTACTCCGGCGCA 720
QY 767 cggcgcgagcgggtctccagcgtaacgcgaacaacatcccaagtgaccggaagact 826
DB 721 CGCGCGCAGCGGCTTCCAGCGTACGCGACAAACAAATCCCCAAGTGGACCGGAAGACT 780
QY 827 ggaagatcggttccaaactgtgcaaccagaacaaatcagactgtttctacagacataact 886
DB 781 GGAAGATCGGCTTCCAACTGTGCAACCAAGCAAAATCAGACTGTTCTACCGACATACT 840
QY 887 cctctggggtgtagtcagtgagggagtggtacgcgttccattacatacaacattctgtcca 946
DB 841 CCTCTGGGTGGATGCAGTGAGGGAGTGTACCGCTTCCATTACATCAACATTTCTCCA 900
QY 947 gactgtcgaacacctgcgcgtctctagagaagaagacccctgggcaactcattcact 1006
DB 901 GACTGTGCGACACCTCGCCCGCTCTAGAGGAAGAGCCCTGGGCAACTTCATCTTCACT 960
QY 1007 gtctgtcaaccagggccctgcgaaccagcgcaatcttccaaagttccaccaccccatgt 1066
DB 961 GTCGCTTCAACCAAGGCCCTGCAACCAAGCGGAATTTATCCRAAGTTCCACCAACCCATGT 1020
QY 1067 acgggaactgctacacttccaatgacaagaacaactcccaatctctgtgactcctcagc 1126
DB 1021 ACGGGAACCTGCTACACTTTCAATGACAAGAAACAATCCAAATCTCTGGATGTCTCCATGC 1080
QY 1127 ctggagtcacaagtgttctccctgacactgcgcagacagacagaatgacttcatcccc 1186
DB 1081 CTGGAGTCAACAATGTTGTCCCTGACACTGCGCACAGACAGAGAATGACTTTCATCCCC 1140
QY 1187 tgctgtccacagtgacggggggccaggggtgatggtgcagtgcaggtgagcctgaccttta 1246
DB 1141 TGCTGTCCACAGTGACGGGGCCAGGGTGATGTTGTCATGGTGACAGATGAGCCTGCTTTA 1200
QY 1247 tggatgattggtggttcaacttgagcctggcgtggagacctccatcagtatgagaaagg 1306
DB 1201 TGGATGATGGTGGCTTCAACTTGAGGCCTTGGCGTGAGAGCTCCATCAGTATGAGAAAG 1260
QY 1307 aagccctgacagcctcgagagaataacggcactgtactgagaatggtagcagtgctc 1366
DB 1261 AAGCCCTGACACGCTCGGAGGAAATTAGCGGACCTGTACTGAGAATGGTAGCGATGTC 1320
QY 1367 csgtcaagaacctttacccttccaagtatacacagcaggtgcttcactcctcctgttcc 1426
DB 1321 CGGTCGAAGAACCTTTTACCCTTCAAGTATACAGCAGGGTGTGATTCATCTCTGCTTCC 1380
QY 1427 aggagaacatgatcaagaagtgtggctgtgctacatcttctaccctaaagcccaaggag 1486
DB 1381 AGGAGAACATGATCAGAGAGTGTGGCTGTGCCCTACATCTTCTACCTTAAGCCCAAGGAG 1440
QY 1487 ttgagttctgtactaccgaagcagagctcctggggctattgctattataaactcagg 1546
DB 1441 TTGAGTTCTGTGACTACCCGAAAGCAGAGCTCTCGGGCTATTGCTATTATAAACTGCAGG 1500

QY 1547 ggccttctcttgagcagcctgggctgtttctccaagtcggaagcctttagtgtga 1606
DB 1501 GCGCCTTCTCTTGGAGAGCCTGGGCTGTTTCTCCAAGTCTCGGAAGCCTTGTAGTGTGA 1560
QY 1607 tcaactacaactctctgcgcggtactcaacggtggccatctgtgaagtcaccagattgga 1666
DB 1561 TCAACTACAAACTCTCTGCCGCTACTCACGGTGGCCATCTGTAAAGTCCAGGATTGGA 1620
QY 1667 tcttcagatgctgtccttgcagacaattacactattaaacacaaagaaacggagttg 1726
DB 1621 TCTTCCAGATGCTGTCTTGCAGAACAAATTACACTATTAAACACAAAGAAACGGAGTTG 1680
QY 1727 caaagctcaacatcttcttcaagagctgaactataaaactaatctcgagcttctcctctg 1786
DB 1681 CAAAGCTCAACTCTTCTTCAAGGAGTGAACATATAAACTAATTTCGGAGTCTCCTTCTG 1740
QY 1787 tcaagatggtcaacctctctgtccaaacctggggcagccaagtggagcctgtggttggctcgt 1846
DB 1741 TCACGATGGTTCAGCCTCTCTTCCAACCTGGGCGAGCCAGTAGTGAGCCTGTGTTTGGCTCGT 1800
QY 1847 cgtgtctctgtggtggagatgacgagactcattcttcgacctcctggtcatcacacttc 1906
DB 1801 CCGTGTCTCTGTGTGGAGATGGCGAGCTCATCTTCGACCTCCTGTGCTCATCACACTTC 1860
QY 1907 tcatgtgtctacgccggttccggagccggttactggtctccagagacgagggcgaggggtg 1966
DB 1861 TCATGTGTCTACGCCGTTTCCGGAGCCGGTACTGGTCTCCAGGACGAGGGGCCAGGGGTG 1920
QY 1967 ccaggaggtggcctccaaactccctcctcctccogtcccgctttctgtctcctcacta 2026
DB 1921 CCAGGAGGTGGCTCCACTCCAGTCTCCTTCCCTCCGTCCTGCTTCTGCTCACCCCTA 1980
QY 2027 catcccaacacacttcttggcccgagcagggcaatgacctccctcctggcctgacagccc 2086
DB 1981 CATCCCAACACCTTCTTGGCCAGCAGGGCATGACCCCTCCCTTCCCTTGACAGCCC 2040
QY 2087 ctcaactgctactgacttagggcccgagtgccctcactggagactctgcggcgctg 2146
DB 2041 CTCACCTGCTATGCTACTCTAGGCCCCAGTGCCCTCCACTGGACTCTGCGGCGCTG 2100
QY 2147 actgtctgctgtgcctggcgctctgtgagaggagagagatcctctcaccagcc 2206
DB 2101 ACTGTTCTGTGCTGTGCCCTGGCGCGCTCTGAGAGAGAGAGGATCCTCTCACCCAGCC 2160
QY 2207 ctgagctccttgaactgatgatctcacttcagca 2246
DB 2161 CTGAGCTCCCTTGAACATGATGATATCTCACTTTCAGCA 2200

RESULT 4

RNU54699 2200 bp mRNA ROD 14-JUN-1997
LOCUS Rattus norvegicus epithelial sodium channel alpha subunit (rEnaca)
DEFINITION mRNA, complete cds.
ACCESSION U54699
VERSION U54699.1 GI:2148925
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2200)
AUTHORS Kreutz,R., Struk,B., Rubattu,S., Hubner,N., Szpirer,J., Szpirer,C.,
Ganten,D. and Lindpaintner,K.
TITLE Role of the alpha-, beta-, and gamma-subunits of epithelial sodium
channel in a model of polygenic hypertension
JOURNAL Hypertension 29 (1), 131-136 (1997)
MEDLINE 97191134
AUTHORS 2 (bases 1 to 2200)
REFERENCE Kreutz,R.
AUTHORS Direct Submission
TITLE Submitted (11-APR-1996) Reinhold Kreutz, Medicine, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115,
JOURNAL


```
QY 1667 tcttcgagatgctgtctcttcgagacaatactactattaaacaagaacgagtg 1726
Db 1621 TCCTCGAGATGCTGTCTCTTCAGAACAAATACACTATTAAACAACAAAGACGAGTTG 1680
QY 1727 caaagctcaacatctcttcctcaagagctgaactataaaactaatcctgagctctctg 1786
Db 1681 CAAAGCTCAACATCTTCTTCAAGGAGCTGAACATATAAAACATAATTCGGAGTCTCTTCTG 1740
QY 1787 tccgatggctcagctctcttccaaactggcagcagtgagcctgtgttggctcgt 1846
Db 1741 TCAGATGTCACGCTCTCTGTCACCTGGCAGCCAGTGGAGCTGTGTTGGCTCGT 1800
QY 1847 ccgtgctctctgtgtggagatgctgacgacgtcatcttcctgacctctgtgcatcacattc 1906
Db 1801 CCCTGCTCTCTGTGTGGAGATGGCGAGCTCATCTTCGACCTCCTGTGTCATCACACTTC 1860
QY 1907 tcatgtgtctacccgggttcctgagccgggtactggtctccagagcagggggcagggggtg 1966
Db 1861 TCATGCTGCTACCGCGGTTCGAGCGGTACTGCTCTCCAGGACGAGGGGCGAGGGGTG 1920
QY 1967 ccaggaggtggctccactccagcttctctctccctccctccgttctctgtctccacctta 2026
Db 1921 CCAGGAGGTGGCTCTCCACCTCCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
QY 2027 catcccccacacctctcttggcccccagcaggcatgacacctccctcgccctgacagccc 2086
Db 1981 CATCCACACACCTTCTTTCGCCCCAGCAGGGCATGCCCTCCCTGGCCCTGCACGCC 2040
QY 2087 ctccacctgctatgctactctctaggcccccagtgccctccactggactctgcggcgctg 2146
Db 2041 CTCACCTGCTATGCTACTCTTAGGCCCCAGTGCCTCTCCACTGGACTCTGCGGCGCTG 2100
QY 2147 actgtctcctgtgcccctgagcgtctgagagagagaagatcctctcaccagcc 2206
Db 2101 ACTGTTCTGCTGTGCTGCGCGCGCTCTGAGAGAGAGAAGATCTCTCACCCAGCC 2160
QY 2207 ctgagctccttgtaactgatgatgatctcacttcagca 2246
Db 2161 CTGAGCTCCCTTGTAACTGATGATATCTCACTTCACGA 2200

RESULT 5
AF082073 749 bp mRNA ROD 03-MAR-1999
LOCUS Rattus norvegicus amiloride-sensitive epithelial sodium channel
DEFINITION alpha subunit (EnaCA) mRNA, partial cds.
ACCESSION AF082073
VERSION AF082073.1 GI:4322357
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 749)
AUTHORS Otulakowski,G., Rafii,B., Bremner,H.R. and O'Brodoovich,H.
TITLE Structure and Hormone Responsiveness of the Gene Encoding the
alpha-Subunit of the Rat Amiloride-sensitive Epithelial Sodium
Channel
JOURNAL Am. J. Respir. Cell Mol. Biol. (1999) In press
REFERENCE 2 (bases 1 to 749)
AUTHORS Otulakowski,G., Rafii,B., Bremner,H.R. and O'Brodoovich,H.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-1998) Lung Biology, Hospital for Sick Children,
555 University Avenue, Toronto, Ont M5G 1S3, Canada
FEATURES
source Location/Qualifiers
1..749
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/note="5' RACE product"
1..518
/gene="EnaCA"
1..749
```

```
BASE COUNT 194 a 203 c 263 g 89 t
ORIGIN
Query Match 9.1%; Score 283; DB 12; Length 749;
Best Local Similarity 100.0%; Pred. NO. 8.9e-156;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 27 cagcctgggagtcgggcacggtcccgacagccccattctcttcacgcgtaatgatgct 86
Db 467 CAGCTGGGATCGGGCAGCGTCCGGCAGACGCCCATCTGCTTCAGCTAATGATGCT 526
QY 87 ggaccacacagagccccctgagctcaacattgacattgacacctcacgcctccaactgcc 146
Db 527 GGACCACACAGAGCCCTGAGCTCAACATTGACCTTAGACCTTCCAGCGCTCCAACTGCC 586
QY 147 taaggggtccatgaagggaacacaaattcaaggagagcaagacctgtctctcagcccat 206
Db 587 TAAGGGGTCCATGAAGGGCAACCAATTCAAGGAGCAAGACCTTGTCTCTCAGCCCAT 646
QY 207 gcaaggactggggaagggggacaaacgtgaagagcagggcctggggcggaacctcagc 266
Db 647 GCAGGACTGGGAAGGGGACAAACGTGAAGACAGAGGGCTGGGCGGGAACCTCAGC 706
QY 267 accccggcagccaccaggaggaggaggagcactgattgaattc 309
Db 707 ACCCGGCGAGCCACCAGGAGGAGGAGGAGGCACTGATTGAATTC 749

RESULT 6
AF081783 2223 bp DNA ROD 03-MAR-1999
LOCUS Rattus norvegicus amiloride-sensitive epithelial sodium channel
DEFINITION alpha subunit (EnaCA) gene, partial cds.
ACCESSION AF081783
VERSION AF081783.1 GI:4322328
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2223)
AUTHORS Otulakowski,G., Rafii,B., Bremner,H.R. and O'Brodoovich,H.
TITLE Structure and hormone responsiveness of the gene encoding the alpha
subunit of the rat amiloride-sensitive epithelial sodium channel
JOURNAL Am. J. Respir. Cell Mol. Biol. (1999) In press
REFERENCE 2 (bases 1 to 2223)
AUTHORS Otulakowski,G., Rafii,B., Bremner,H.R. and O'Brodoovich,H.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-1998) Lung Biology, Hospital for Sick Children,
555 University Avenue, Toronto, Ont M5G 1S3, Canada
FEATURES
source Location/Qualifiers
1..2223
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
1478..2223
/gene="EnaCA"
/note="alternative transcription start site"
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
1478..2223
/gene="EnaCA"
```

```
mRNA      1931. .>2223
/genes="ENACA"
/notes="alternative transcription start site"
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
1993. .>2223
/genes="ENACA"
/notes="ion channel"
/codon_start=1
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
/protein_id="AAB16017.1"
/db_xref="GI:4322329"
/translation="MMLDHTRAPELNLDLHASNSPKGSMKGNQKPEDPCPPQMQ
GLKGDKREEQGLGPPSPAPROPTEEALIEP"

BASE COUNT      432 a      673 c      639 g      479 t
ORIGIN

Query Match      9.1%; Score 283; DB 12; Length 2223;
Best Local Similarity 100.0%; Pred. No. 9.9e-156;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 cagctggatgoggcagcgtcccgacagcccccattctgccttcacgtatgatgct 86
|||||
Db 1941 CAGCCTGGGATGCGGGCAGCGTCCCGGACAGCCCATTCGCTTCACGGCTAATGATGCT 2000
|||||

QY 87 ggaccacacagacccttgatgctcaacattgaactgaacttcaacgctcccaactgcgc 146
|||||
Db 2001 GGACCACACAGAGCCCTCAGCTCAACATTTGACCTAGACCTTCACGCTCCCAACTCGCC 2060
|||||

QY 147 taagggtccatgaaggcgaacaaattcaaggagcagccctgtctcctcagcccat 206
|||||
Db 2061 TAAGGGTCCATGAAGGGCAACCAATTCAGGACCAAGACCCTTGTCTCTCAGCCCAT 2120
|||||

QY 207 gcaaggactgggaagggggacaaactgaagagcagggcctggcccggaacctcagc 266
|||||
Db 2121 GCAAGGACTGGGAAGGGGACAAACCTGAAGACGAGCGCTGGCCCGGAACCTCAGC 2180
|||||

QY 267 accccgcagccaccagagagagaggaggagcactgattgaattc 309
|||||
Db 2181 ACCCCGCAGCCACCAGGAGGAGGAGGAGGCACTGATTGAATTC 2223
|||||

RESULT 7
RNC DK3UTR      252 bp      mRNA      ROD      04-APR-1997
LOCUS      R.norvegicus mRNA for 3'UTR of amiloride sensitive sodium channel
DEFINITION      protein.
ACCESSION      X89818
VERSION      X89818.1 GI:1419186
KEYWORDS      3' UTR; amiloride sensitive sodium channel protein; CDK3 protein.
SOURCE      Norway rat.
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE      1 (bases 1 to 252)
AUTHORS      Page,R., Morris,C., Williams,J., von Ruhland,C. and Malik,A.N.
TITLE      Isolation of diabetes-associated kidney genes using differential
display
JOURNAL      Biochem. Biophys. Res. Commun. 232 (1), 49-53 (1997)
MEDLINE      97236278
REFERENCE      2 (bases 1 to 252)
AUTHORS      Page,R.A.
TITLE      Direct Submission
JOURNAL      Submitted (18-JUL-1995) R.A. Page, University of Wales College of
Cardiff, School of Molecular and Med.Biosciences, P.O. Box 911,
Cardiff CF1 3US, UK
COMMENT      Related sequence X70521.
FEATURES      Location/Qualifiers
source      1. .252
/organism="Rattus norvegicus"
/strain="Goto-Kakizaki (GK) Wistar"

db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="kidney"
/clone="CDK3 (pSEM clone)"
1. .252
/genes="CDK3"
1. .252
/genes="CDK3"
/notes="amiloride sensitive sodium channel protein"
221. .226
/polyA_signal
/genes="CDK3"
231
/polyA_site
/genes="CDK3"
231
BASE COUNT      48 a      77 c      51 g      76 t
ORIGIN

Query Match      4.5%; Score 141; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.3e-72;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2905 ccctcacccccacagtgactgtattatattatgttcacattctttgtgtcctccctcgtg 2964
|||||
Db 62 CCCTCACCCCCACAGTACTGTATTATATGTTTCACATTTCTTTGTGTCTCCCTCCCTGG 121
|||||

QY 2965 gtaaacctgagctcctgtgtggtcagggatgagattgctctgtttgtatcctccggt 3024
|||||
Db 122 GTAAACCTGAGCTCCTGTGTGTGTCAGGGATGAGATTGCTCTGTTTGTATCTTCCGGT 181
|||||

QY 3025 ctagcccagctcccaacttg 3045
|||||
Db 182 CTAGCCCAGTCTCCCACTGG 202
|||||

RESULT 8
AF112185      3000 bp      mRNA      ROD      27-MAR-2000
LOCUS      Mus musculus epithelial sodium channel alpha subunit mRNA, complete
DEFINITION      cds.
ACCESSION      AF112185
VERSION      AF112185.1 GI:4469398
KEYWORDS
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 3000)
AUTHORS      Ahn,Y.J., Brooker,D.R., Kosari,F., Harte,B.J., Li,J., Mackler,S.A.
and Kleyman,T.R.
TITLE      Cloning and functional expression of the mouse epithelial sodium
channel
JOURNAL      Am. J. Physiol. 277 (1), F121-F129 (1999)
MEDLINE      99345762
REFERENCE      2 (bases 1 to 3000)
AUTHORS      Ahn,Y.J., Brooker,D.B. and Kleyman,T.R.
TITLE      Direct Submission
JOURNAL      Submitted (09-DEC-1998) Renal Electrolyte Division, University of
Pennsylvania, 422 Curie Boulevard, 700 Clinical Research Building,
Philadelphia, PA 19104, USA
FEATURES      Location/Qualifiers
source      1. .3000
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="6"
/map="distal"
32. .2131
/codon_start=1
/product="epithelial sodium channel alpha subunit"
/protein_id="AAB21244.1"
/db_xref="GI:4469398"
/translation="MLDHTRAPELNLDLHASNSPKGSMKGNKFNKEDLCPPPLPMQG
LGKDKREQALGPPEPQPTPEEALIEFHSYRELFFQFCNNNTTIHGAIRLVC
```

KNNRMTAFMAVWLMLCTFGMYWQFALLEEYFSPVSLNINLNSDKLVFPVAVTVCTL
NRYRYTEIKEDLEEDRIPEQTLFDLYKYNSTYRQAGGRRSTRDLRGALPHQLRL
RTPPPNPARSARSASSVRDNNPOVDKDWIKGFCOLCNONKSDCFYQYSSGVDAVR
EWKRYHINILSRPDTSPALKEEALGSFIFTCRNOAPCNOANTSOFHHPMYGNCYT
FNKNLNLMWSMFGVNNGLSLTLRTEONDIRPILSTVTGKRVVHQODEPAPMDGG
GNVRPGVETISIMRKEALDSLGGNYGDCTEGSDVPVKNLYPSKYTQOVCHSCFQE
NNIKKCGCAIIFYPKPKQGFCDYLKQSSWGVCYYKLOAFSLDSLGCFSKCKPCSV
TNYKLSAGYSRWPVSQDWIFEMLSLQNNYTIINNRNGVAKLINIFFKELNKTNSCS
PSVTWVSLNSLGSOWSLWFGSSVLSVEMAEILFDLLVITLIMLLHRRFSRWSPGR
GARGAREVASTPASSPSPRCPHPTSPPSLSLQQCTTPPLALTAPPAYATLGPSP
LDSAVPGSSACAPANAL"

BASE COUNT 666 a 958 c 701 g 675 t
ORIGIN

Query Match 3.3%; Score 104; DB 12; Length 3000;
Best Local Similarity 100.0%; Pred. No. 9.6e-50;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1120 tccatgcctggagtcacaaatggtttgtccctgacactgcgcacagacagaatgacttc 1179
|||||

Db 1070 TCCATGCCTGGACTCAACATGTTGTCCCTGACACTGCGCAGACAGAGATGACTTC 1129
|||||

QY 1180 atccccctgctccacagtcagcggggccagggtgatgtgca 1223
|||||

Db 1130 ATCCCCCTGCTCCACAGTACGGGGGCCAGGCTGATGTGCA 1173
|||||

RESULT 9

I78503 98 bp DNA PAT 03-APR-1998
LOCUS
DEFINITION Sequence 14 from patent US 5693756.

ACCESSION I78503

VERSION I78503.1 GI:3014557

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 98)

AUTHORS Li,X., Blackshaw,S. and Snyder,S.H.

TITLE Amiloride-sensitive sodium channel and method of identifying substances which stimulate or block salty taste perception

JOURNAL Patent: US 5693756-A 14 02-DEC-1997;

FEATURES Location/Qualifiers

1..98

/organism="unknown"

BASE COUNT 18 a 25 c 26 g 29 t

ORIGIN

Query Match 3.1%; Score 98; DB 5; Length 98;

Best Local Similarity 100.0%; Pred. No. 2.4e-46;

Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1517 cctgggctattgtctattataaactcagggcgccctctctctggacagcctggctgtt 1576
|||||

Db 1 CCTGGGCTATTGCTATTATAACTCGAGGCGCCTCTCTTGGACAGCCTGGGCTGTT 60
|||||

QY 1577 tctccaagtgcgaagcctttagtgcataactac 1614
|||||

Db 61 TCTCCAAGTGTGGAAGCCTTGTAGTGTGATCACTAC 98
|||||

RESULT 10

AF002665

LOCUS

DEFINITION Rattus norvegicus epithelial sodium channel alpha subunit (Scnnla)

ACCESSION AF002665

VERSION AF002665.1 GI:4100903

KEYWORDS

SOURCE Norway rat.

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 852)
AUTHORS Stock,P., Kreutz,R., Ganten,D. and Lindpaintner,K.

TITLE Identification of two microsatellites in the gene coding for the alpha epithelium sodium channel subunit in the rat

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 852)

AUTHORS Kreutz,R.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-1997) Clinical Pharmacology, Benjamin Franklin Hospital, Free University of Berlin, Hindenburgdamm 30, Berlin D-12200, Germany

FEATURES Location/Qualifiers

1..852

/organism="Rattus norvegicus"

/strain="Stroke-prone spontaneously hypertensive rat (SHRSPHD), Heidelberg strain"

/db_xref="taxon:10116"

/chromosome="4"

/map="4q42: Spr-D4Mit20-Eno2-Scnnla"

/clone="pCRII-8-Scnnla"

<1..>114

/gene="Scnnla"

/product="epithelial sodium channel alpha subunit"

<1..>114

/gene="Scnnla"

<1..>114

/note="RNASNAAC; RNENACA; similar to epithelial sodium channel alpha subunits encoded by GenBank Accession Numbers X70497 and X70521"

/codon_start=1

/product="epithelial sodium channel alpha subunit"

/protein_id="AAD01004.1"

/db_xref="GI:4100904"

/translation="LRTPPPPYSGRTARGSSVRDNNPQVDKRWKIGQL"

BASE COUNT 202 a 200 c 253 g 197 t

ORIGIN

Query Match 2.3%; Score 71; DB 12; Length 852;

Best Local Similarity 100.0%; Pred. No. 2.9e-30;

Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 776 gcgggtcttcagtcgcgcacacacatcccccaagtcgaccggaagactggaagatcg 835
|||||

Db 44 GCGGGTCTTCAGGTACGCGACACAAATCCCCAAGTGGACCGGAGGACTGGAAGATCG 103
|||||

QY 836 gcttccaactg 846
|||||

Db 104 GCTTCCAACATG 114
|||||

RESULT 11

CPO249296

LOCUS

DEFINITION Cavia porcellus mRNA for sodium channel, alpha subunit.

ACCESSION AJ249296

VERSION AJ249296.1 GI:6687302

KEYWORDS alpha subunit; sodium channel.

SOURCE domestic guinea pig.

ORGANISM Cavia porcellus

REFERENCE 1 (bases 1 to 2695)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Cavidae; Cavia.

Schnizler,M., Mastroberardino,L., Reifarth,F., Weber,W.M., Verrey,F. and Claus.W.

TITLE CAMP sensitivity conferred to the epithelial Na+ channel by alpha-subunit cloned from guinea-pig colon

JOURNAL Pflugers Arch. 439 (5), 579-587 (2000)

MEDLINE 2025288

REFERENCE 2 (bases 1 to 2695)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 852)

AUTHORS Stock,P., Kreutz,R., Ganten,D. and Lindpaintner,K.

TITLE Identification of two microsatellites in the gene coding for the alpha epithelium sodium channel subunit in the rat

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 852)

AUTHORS Kreutz,R.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-1997) Clinical Pharmacology, Benjamin Franklin Hospital, Free University of Berlin, Hindenburgdamm 30, Berlin D-12200, Germany

FEATURES Location/Qualifiers

1..852

/organism="Rattus norvegicus"

/strain="Stroke-prone spontaneously hypertensive rat (SHRSPHD), Heidelberg strain"

/db_xref="taxon:10116"

/chromosome="4"

/map="4q42: Spr-D4Mit20-Eno2-Scnnla"

/clone="pCRII-8-Scnnla"

<1..>114

/gene="Scnnla"

/product="epithelial sodium channel alpha subunit"

<1..>114

/gene="Scnnla"

<1..>114

/note="RNASNAAC; RNENACA; similar to epithelial sodium channel alpha subunits encoded by GenBank Accession Numbers X70497 and X70521"

/codon_start=1

/product="epithelial sodium channel alpha subunit"

/protein_id="AAD01004.1"

/db_xref="GI:4100904"

/translation="LRTPPPPYSGRTARGSSVRDNNPQVDKRWKIGQL"

BASE COUNT 202 a 200 c 253 g 197 t

ORIGIN

Query Match 2.3%; Score 71; DB 12; Length 852;

Best Local Similarity 100.0%; Pred. No. 2.9e-30;

Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 776 gcgggtcttcagtcgcgcacacacatcccccaagtcgaccggaagactggaagatcg 835
|||||

Db 44 GCGGGTCTTCAGGTACGCGACACAAATCCCCAAGTGGACCGGAGGACTGGAAGATCG 103
|||||

QY 836 gcttccaactg 846
|||||

Db 104 GCTTCCAACATG 114
|||||

RESULT 11

CPO249296

LOCUS

DEFINITION Cavia porcellus mRNA for sodium channel, alpha subunit.

ACCESSION AJ249296

VERSION AJ249296.1 GI:6687302

KEYWORDS alpha subunit; sodium channel.

SOURCE domestic guinea pig.

ORGANISM Cavia porcellus

REFERENCE 1 (bases 1 to 2695)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Cavidae; Cavia.

Schnizler,M., Mastroberardino,L., Reifarth,F., Weber,W.M., Verrey,F. and Claus.W.

TITLE CAMP sensitivity conferred to the epithelial Na+ channel by alpha-subunit cloned from guinea-pig colon

JOURNAL Pflugers Arch. 439 (5), 579-587 (2000)

MEDLINE 2025288

REFERENCE 2 (bases 1 to 2695)

AUTHORS Schnizler, M.K.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Schnizler M.K., Biologie, Institut fuer
tierphysiologie, Justus-Liebig-Universitaet, Wartweg 95, 35392
Giessen, GERMANY

FEATURES source
1. .2695
/organism="Cavia porcellus"
/db_xref="taxon:10141"
/country="Germany"
/sex="male"
/dev_stage="adult"
/tissue_type="colonic mucosa"
/tissue_lib="colonic mucosa"
293. .2263
/function="epithelial sodium channel"
/codon_start=1
/product="alpha subunit"
/protein_id="CAB64910.1"
/db_xref="GI:6687303"
/translations="MKGDELKEOGLPPPOPLQGLKCEQPLGPEPTAPOOHTTEE
EEALTEFHSYRELFOFFCNNTIIGHAIRLVCSKHNRMKTAFAWVLMCTFGMYWQF
ALLGFEYSFVSLNINLNSDKLVFAVTVCTLNRYKEIKQLRELDRIQTOTLFD
LYNNASSTLLAGARRSLADTLPLQRPVQPEPRARRSDPSVRDNNRVDNR
DMRVGQLCNQKSDCFYQTSSEGVGVREWFHYINILAQVADTSPSLEELALGNF
IFACRNOAPCTOENTSHPHPIYGCYTNENKNDSSLWASMPGINGLSLTLRTQ
NDYIPLLSVTGARVTVHGDDEPAFMDGDFNLRPGVETISMRKEALDRGSGYDC
TQGSVPQVNLPSKYTQVCIHSCFOENMKQCCAYIFPKPGVQPCDTRKQSA
WGYTYKLGAFSSDLSKCNKPCNVITYIKLSAGYSRWPSTVSDMTFQWLSLN
NYTISKRNGVANLYFKELNYRTNSESPTVMVLLSLNGQSLWFGSSVLSVYE
MAFMFDLLVLTLLMLRFRSRYWSPGRGARAAREVACTPPPSLPSRFCAHSAFPTL
TAPPPAYATLSACPPLQGLAGASSAACAPREP"

CDS
1. .2695
/organism="Cavia porcellus"
/db_xref="taxon:10141"
/country="Germany"
/sex="male"
/dev_stage="adult"
/tissue_type="colonic mucosa"
/tissue_lib="colonic mucosa"
293. .2263
/function="epithelial sodium channel"
/codon_start=1
/product="alpha subunit"
/protein_id="CAB64910.1"
/db_xref="GI:6687303"

Query Match 1.9%; Score 59; DB 12; Length 2695;
Best Local Similarity 100.0%; Pred. No. 4.2e-23;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1786 gtacagatggtcagctctctgacactgcgcacagcagaatgactt 1844
|||||
Db 1910 GTCACGATGTCAGCCCTCTCCAACTGGCGAGCCAGTGTGTTGGCTC 1968
|||||

RESULT 12
I78500/c
LOCUS 178500 45 bp DNA PAT 03-APR-1998
DEFINITION Sequence 11 from patent US 5693756.
ACCESSION I78500
VERSION I78500.1 GI:3014654
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Li, X., Blackshaw, S. and Snyder, S.H.
TITLE Amiloride-sensitive sodium channel and method of identifying
substances which stimulate or block salty taste perception
JOURNAL Patent: US 5693756-A 11 02-DEC-1997;
FEATURES source
1. .45
/organism="unknown"
BASE COUNT 10 a 10 c 12 g 13 t
ORIGIN

Query Match 1.9%; Score 59; DB 12; Length 2695;
Best Local Similarity 100.0%; Pred. No. 4.2e-23;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1786 gtacagatggtcagctctctgacactgcgcacagcagaatgactt 1844
|||||
Db 1910 GTCACGATGTCAGCCCTCTCCAACTGGCGAGCCAGTGTGTTGGCTC 1968
|||||

RESULT 12
I78500/c
LOCUS 178500 45 bp DNA PAT 03-APR-1998
DEFINITION Sequence 11 from patent US 5693756.
ACCESSION I78500
VERSION I78500.1 GI:3014654
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Li, X., Blackshaw, S. and Snyder, S.H.
TITLE Amiloride-sensitive sodium channel and method of identifying
substances which stimulate or block salty taste perception
JOURNAL Patent: US 5693756-A 11 02-DEC-1997;
FEATURES source
1. .45
/organism="unknown"
BASE COUNT 10 a 10 c 12 g 13 t
ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 5.4e-15;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1134 caacaatggtttccctgacactgcgcacagcagaatgactt 1178
|||||

Db 45 CAACAATGTTTTCCTGACACTGCGCACAGCAGAGTAATGACTT 1

RESULT 13
I78501/c
LOCUS 178501 45 bp DNA PAT 03-APR-1998
DEFINITION Sequence 12 from patent US 5693756.
ACCESSION I78501
VERSION I78501.1 GI:3014655
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Li, X., Blackshaw, S. and Snyder, S.H.
TITLE Amiloride-sensitive sodium channel and method of identifying
substances which stimulate or block salty taste perception
JOURNAL Patent: US 5693756-A 12 02-DEC-1997;
FEATURES source
1. .45
/organism="unknown"
BASE COUNT 11 a 11 c 10 g 13 t
ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 5.4e-15;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1498 gactaccgaagcagagctctctggtgctattgctattataaactg 1542
|||||

Db 45 GACTACGAAAGCAGAGCTCTCTGGGCTATTGCTATTATAAATG 1
|||||

RESULT 14
OCU132108
LOCUS OCU132108 2915 bp mRNA MAM 13-JAN-1999
DEFINITION Oryctolagus cuniculus mRNA for epithelial sodium channel, alpha
subunit.
ACCESSION AJ132108
VERSION AJ132108.1 GI:4158223
KEYWORDS alpha subunit; alpha-ENaC gene; epithelial sodium channel.
SOURCE Oryctolagus cuniculus.
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 2915)
AUTHORS Kudlacek, O.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1999) Kudlacek O., Pharmacological Institute,
University of Vienna, Waehringerstrasse 13A, A-1090 Wien, AUSTRIA
REFERENCE 2 (bases 1 to 2915)
AUTHORS Kudlacek, O., Weisz, E., Wiener, H. and Plass, H.
TITLE The rabbit epithelial sodium channel
JOURNAL Unpublished
FEATURES source
1. .2915
/organism="Oryctolagus cuniculus"
/db_xref="taxon:9986"
35. .1957
/gene="alpha-ENaC"
35. .1957
/gene="alpha-ENaC"
/codon_start=1
/product="epithelial sodium channel, alpha subunit"
/protein_id="CAA10571.1"
/db_xref="GI:4158224"

Translation="MKGDKREOGPGPETVAPQOQTEDEALIEFHSYRELFOFFCN
NTIIGHAIRLVCSKHNRMKTAFAWVLMCTFGMYWQFGLLFGFYFSPYNLNLS
DKLVFPVAVTVCTLNIPRYPEITQELKELSDITQOTLLDLFKYNASTLEAQPHRRDVH
PLPHLPQRLRVPPVPRLEARRASSSDNSPEVGRKQMMIGFQLCNQNRSDCFYQ
RYSGGVDAVREYRFHYINILSRSLDTSLSRQLGNFTTCRFNQAFCGDNGYSHFHH
PMYGNCYTFNDKNNSSLWMSMPCINNGLSLTLRTQNDFTIPLLSTVTGARVMVHGQD

EPAFMDGGFNLRPGVETISMRKESLDRLGGDYGDCTQNGSDVPVKNLYRSKYTQOV
CIHSPDMYKCGCAYIFYPLPEGYCDYRKHNGWCYCYKLDQAFSSDRLGCEFT
KCRKPCSVNTYELSAGYSRPSVTSDQVYFOMLSQNNYTVSNKRNGVAKNIYFKEL
NRYKANSESPVMTYLLSLNLSGMSLAFSGSVLSVEMAEILFDLSVITFTMLLRFR
SRWSPGRGAGREAVASPSVALPSRCPHTSPSPQPGFTLPPSLATPPAYATL
GPCLSQSGSACAPGP"
BASE COUNT 577 a 978 c 774 g 586 t
ORIGIN

Query Match 1.3%; Score 41; DB 3; Length 2915;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 gagctctccagttcttgcacaacaccaccatccacgg 365
|||||
Db 143 GAGCTCTCCAGTCTTCTGCAACACACCACCATCCACGG 183
|||||

RESULT 15
PAU95179 636 bp mRNA PLN 15-DEC-1997
LOCUS Prunus armeniaca putative Nifu protein mRNA, partial cds.
DEFINITION
ACCESSION U95179
VERSION U95179.1 GI:2688825
KEYWORDS
SOURCE apricot.
ORGANISM Prunus armeniaca

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Rosales; Rosaceae; Prunus.
REFERENCE 1 (bases 1 to 636)
AUTHORS Mbequie A Mbequie,D., Gomez,R.-M. and Fills-Lycaon,B.
TITLE Molecular cloning and nucleotide sequence of a putative Nifu
protein from apricot
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 636)
AUTHORS Mbequie A Mbequie,D., Gomez,R.-M. and Fills-Lycaon,B.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-1997) Station de Technologie des Produits
Vegetaux, INRA, Site AGROPARC, domaine Saint Paul, Avignon 84914
cedex 9, France

FEATURES
source
1. .636
/organism="Prunus armeniaca"
/strain="bergeron"
/db_xref="taxon:36596"
/tissue_type="mesocarp; exocarp"
/dev_stage="ripe fruit"
/clone="pAPRI31"
<1. .231
/codon_start=1
/product="putative Nifu protein"
/protein_id="AAB88877.1"
/db_xref="GI:2688826"
/translation="ELAVPNVDVLEDRPYLIADGGDVVSVEDGVVSLKLGACG
SCPSTTTMKGIERVLKEKFGDKLQIVL"
BASE COUNT 199 a 118 c 149 g 170 t
ORIGIN

Query Match 1.3%; Score 40; DB 8; Length 636;
Best Local Similarity 100.0%; Pred. No. 6.4e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 581 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 620
|||||

RESULT 16
GGVITRO 1620 bp mRNA VRT 04-DEC-1997
LOCUS

G.gallus mRNA for vitronectin.
ACCESSION Y11030
VERSION Y11030.1 GI:1922281
KEYWORDS vitronectin.
SOURCE chicken.
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1620)
AUTHORS Martinez-Morales,J.R., Barbas,J.A., Marti,E., Bovolenta,P.,
Egarr,D. and Rodriguez-Rebar,A.

TITLE Vitronectin is expressed in the ventral region of the neural tube
and promotes the differentiation of motor neurons
JOURNAL Development 124 (24), 5139-5147 (1997)
MEDLINE 98090066
REFERENCE 2 (bases 1 to 1620)
AUTHORS Barbas,J.A.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1997) J.A. Barbas, Instituto Cajal CSIC, Av.
Doctor Acre 37, E- 28002 Madrid, SPAIN

FEATURES
source
1. 1620
/organism="Gallus gallus"
/strain="White leghorn"
/db_xref="taxon:9031"
/dev_stage="7 day embryo"
/tissue_type="retina"
/clone_lib="lambda-ZAP II"
141. .1502
/codon_start=1
/product="vitronectin"
/protein_id="CAA71914.1"
/db_xref="GI:1922282"
/translation="MRLLPVLVLLALTRAEDSCGRDEGNAMKKKCCDILCN
YYSCSDSYTVCAKAVTRGDVFPALPEDDYLDYLSIDTGTVPTEAPTEHPTEPH
YSPSPRTVIDATEETPEPEVPTLHPTTTTSDTRNPSLDEPEELCSRKPFN
AFTDLKNGSIYAFKGYELDKSVKPGYKLIQSVGIGPIDAAFRVYVFKGKY
LFKGSYWRFDGALDPCYPRDISEGFIGIPNDIDAAFPALPAHSYHGVRYVFKGKY
YWSDFAHQPTQACEKSSPTVFNHYAFMRDSDWEDIFLSIFGSRMGVAGSORLISR
DWGPNVLDAAAGRIYVSSRQRRSRHRKRYNRHTLNGLGWSLWNSDSSTD
TESWLSGSCQETLQSVYFFVGDKYRYVNLRTKRYDLVQPPYPRISIAQYWLDCQPQDE
EST"
polya_signal 1550. .1555
BASE COUNT 399 a 510 c 431 g 280 t
ORIGIN

Query Match 1.3%; Score 40; DB 4; Length 1620;
Best Local Similarity 100.0%; Pred. No. 7e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 1570 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1609
|||||

RESULT 17
AB002533 2171 bp mRNA PRI 13-FEB-1999
LOCUS Homo sapiens mRNA for Qip1, complete cds.
DEFINITION
ACCESSION AB002533
VERSION AB002533.1 GI:1944124
KEYWORDS Qip1.
SOURCE Homo sapiens cell_line:HeLa S3 cDNA to mRNA.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2171)
AUTHORS Seki,T.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-1997) to the DDBJ/EMBL/GenBank databases.
Takahiko Seki, Tohoku University, Faculty of Pharmaceutical

Sciences; Aoba Aramaki, Aoba-Ku, Sendai, Miyagi 980-77, Japan
(E-mail: takaeghi2.pharm.tohoku.ac.jp, Tel: 81-22-217-6876,
Fax: 81-22-217-6873)
2 (sites)
REFERENCE
AUTHORS Seki, T., Tada, S., Katada, T. and Enomoto, T.
TITLE Cloning of a cDNA encoding a novel importin-alpha homologue, Qip1:
discrimination of Qip1 and Rchl from hsrp1 by their ability to
interact with DNA helixase Q1/RecQ1
JOURNAL Biochem. Biophys. Res. Commun. (1997) In press
FEATURES
source Location/Qualifiers
1. .2171
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HeLa S3"
10. .1575
/gene="Qip1"
10. .1575
/gene="Qip1"
/codon_start=1
/product="Qip1"
/protein_id="BAA19546.1"
/db_xref="GI:1944125"
/translation="MADNEKLDNQRKNKGRDLEMTMRQRNRRVVEVLRKRNKDEH
LKRRNPVEDICEDSDIDGVRQNTSLEATVQNASDNOGQISAVQAARKLLSSD
RNPTDILKSGIPILVHCLERDNPISLOFEAAWLTNIASTQEQAVVQSNVAP
LFLRLHSPHONVCFQAWALGNIIGDGPQCDYVLSGVVKPLLSFISPSIPITFLR
NVTWVNLCRHKDPPPMETIQEILPALCVLIHHTDYNILVTFWALSILNDAGNEQ
IQMVDSIGVPHLVPLSHQEVQTAALRVGNITGTDEQTVVNLCDALSHFPAL
LTHPKRINKAEVFLSNITAGNQOQVADIDNLVPMIHLDDGDFGTQKEAAWAI
SNLTISGRKDVAYLIQONVIPPFCNLLTVKDAQVQVVLGSLNLMKAEDEAETIG
NLIECGGLEKTEQLQNHENEDIYKLAYEIIQDFSSDDIDEDPSLVPEAIGGGTFG
NSSANVPTEGQF"
2171
polyA_site /note="32 A nucleotides"
BASE COUNT 685 a 450 c 477 g 559 t
ORIGIN
Query Match 1.3%; Score 40; DB 9; Length 2171;
Best Local Similarity 100.0%; Pred. No. 7.2e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 1678 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1717
RESULT 18
AC034204
LOCUS AC034204 61001 bp DNA HTG 09-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTB-192J21, WORKING DRAFT SEQUENCE,
16 unordered pieces.
ACCESSION AC034204
VERSION AC034204.2 GI:7528346
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 61001)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 9, 2000 this sequence version replaced gi:7417681.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

-----Summary Statistics
Consensus quality: 47661 bases at least Q40
Consensus quality: 55452 bases at least Q30
Consensus quality: 57169 bases at least Q20
Estimated insert size: 61001; sum-of-contigs estimation
Estimated insert size: 60000; pulse field gel estimation
Quality coverage: 3.20x in Q20 bases; pulse field gel estimation
Quality coverage: 3.46x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1624: contig of 1624 bp in length
* gap of unknown length
* 1625 2807: contig of 1183 bp in length
* gap of unknown length
* 2808 4042: contig of 1235 bp in length
* gap of unknown length
* 4043 5495: contig of 1453 bp in length
* gap of unknown length
* 5496 7142: contig of 1647 bp in length
* gap of unknown length
* 7143 9094: contig of 1952 bp in length
* gap of unknown length
* 9095 10968: contig of 1874 bp in length
* gap of unknown length
* 10969 13357: contig of 2389 bp in length
* gap of unknown length
* 13358 15997: contig of 2640 bp in length
* gap of unknown length
* 15998 20030: contig of 4033 bp in length
* gap of unknown length
* 20031 24641: contig of 4611 bp in length
* gap of unknown length
* 24642 28743: contig of 4102 bp in length
* gap of unknown length
* 28744 33674: contig of 4931 bp in length
* gap of unknown length
* 33675 41268: contig of 7594 bp in length
* gap of unknown length
* 41269 51931: contig of 10663 bp in length
* gap of unknown length
* 51932 61001: contig of 9070 bp in length.
Location/Qualifiers
1. .61001
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTB-192J21"
BASE COUNT 16555 a 14407 c 13948 g 16090 t 1 others
ORIGIN
Query Match 1.3%; Score 40; DB 74; Length 61001;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 31264 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 31303
RESULT 19
AC036192
LOCUS AC036192 80861 bp DNA HTG 07-APR-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-485P2 map 2, LOW-PASS SEQUENCE
SAMPLING.
ACCESSION AC036192

VERSION

AC036192.1 GI:7523861

KEYWORDS

HTGS_PHASE0.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;

AUTHORS

Mammalia: Eutheria: Primates; Catarrhini; Hominidae; Homo.

TITLE

1 (bases 1 to 80861)

JOURNAL

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

REFERENCE

Unpublished

AUTHORS

2 (bases 1 to 80861)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bestien,V., Bieda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Melgrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (07-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9272
Center clone name: 485_P_2

* NOTE: This record contains 89 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 805 904: gap of 100 bp
* 905 1712: contig of 808 bp in length
* 1713 1812: gap of 100 bp
* 1813 2550: contig of 738 bp in length
* 2551 2650: gap of 100 bp
* 2651 3426: contig of 776 bp in length
* 3427 3526: gap of 100 bp
* 3527 4336: contig of 810 bp in length
* 4337 4436: gap of 100 bp
* 4437 5244: contig of 808 bp in length
* 5245 5344: gap of 100 bp
* 5345 6160: contig of 816 bp in length
* 6161 6260: gap of 100 bp
* 6261 7064: contig of 804 bp in length

* 7065 7164: gap of 100 bp
* 7165 7973: contig of 809 bp in length
* 7974 8073: gap of 100 bp
* 8074 8896: contig of 823 bp in length
* 8897 8996: gap of 100 bp
* 8997 9801: contig of 805 bp in length
* 9802 9901: gap of 100 bp
* 9902 10730: contig of 829 bp in length
* 10731 10830: gap of 100 bp
* 10831 11625: contig of 795 bp in length
* 11626 11725: gap of 100 bp
* 11726 12520: contig of 795 bp in length
* 12521 12620: gap of 100 bp
* 12621 13419: contig of 799 bp in length
* 13420 13519: gap of 100 bp
* 13520 14328: contig of 809 bp in length
* 14329 14428: gap of 100 bp
* 14429 15246: contig of 818 bp in length
* 15247 15346: gap of 100 bp
* 15347 16151: contig of 805 bp in length
* 16152 16251: gap of 100 bp
* 16252 17071: contig of 820 bp in length
* 17072 17171: gap of 100 bp
* 17172 17969: contig of 798 bp in length
* 17970 18069: gap of 100 bp
* 18070 18859: contig of 790 bp in length
* 18860 18959: gap of 100 bp
* 18960 19786: contig of 827 bp in length
* 19787 19886: gap of 100 bp
* 19887 20724: contig of 838 bp in length
* 20725 20824: gap of 100 bp
* 20825 21607: contig of 783 bp in length
* 21608 21707: gap of 100 bp
* 21708 22533: contig of 826 bp in length
* 22534 22633: gap of 100 bp
* 22634 23446: contig of 813 bp in length
* 23447 23546: gap of 100 bp
* 23547 24331: contig of 785 bp in length
* 24332 24431: gap of 100 bp
* 24432 25256: contig of 825 bp in length
* 25257 25356: gap of 100 bp
* 25357 26163: contig of 807 bp in length
* 26164 26263: gap of 100 bp
* 26264 27095: contig of 832 bp in length
* 27096 27195: gap of 100 bp
* 27196 28020: contig of 825 bp in length
* 28021 28120: gap of 100 bp
* 28121 28945: contig of 825 bp in length
* 28946 29045: gap of 100 bp
* 29046 29878: contig of 833 bp in length
* 29879 29978: gap of 100 bp
* 29979 30778: contig of 800 bp in length
* 30779 30878: gap of 100 bp
* 30879 31648: contig of 770 bp in length
* 31649 31748: gap of 100 bp
* 31749 32549: contig of 801 bp in length
* 32550 32649: gap of 100 bp
* 32650 33472: contig of 823 bp in length
* 33473 33572: gap of 100 bp
* 33573 34389: contig of 817 bp in length
* 34390 34489: gap of 100 bp
* 34490 35310: contig of 821 bp in length
* 35311 35410: gap of 100 bp
* 35411 36211: contig of 801 bp in length
* 36212 36311: gap of 100 bp
* 36312 37139: contig of 828 bp in length
* 37140 37239: gap of 100 bp
* 37240 38052: contig of 813 bp in length
* 38053 38152: gap of 100 bp
* 38153 38986: contig of 834 bp in length
* 38987 39086: gap of 100 bp
* 39087 39924: contig of 838 bp in length
* 39925 40024: gap of 100 bp

* 40025 40808: contig of 784 bp in length
* 40809 40908: gap of 100 bp
* 40909 41725: contig of 817 bp in length
* 41726 41825: gap of 100 bp
* 41826 42620: contig of 795 bp in length
* 42621 42720: gap of 100 bp
* 42721 43526: contig of 806 bp in length
* 43527 43626: gap of 100 bp
* 43627 44435: contig of 809 bp in length
* 44436 44535: gap of 100 bp
* 44536 45349: contig of 814 bp in length
* 45350 45449: gap of 100 bp
* 45450 46248: contig of 799 bp in length
* 46249 46348: gap of 100 bp
* 46349 47147: contig of 799 bp in length
* 47148 47247: gap of 100 bp
* 47248 48063: contig of 816 bp in length
* 48064 48163: gap of 100 bp
* 48164 48984: contig of 821 bp in length
* 48985 49084: gap of 100 bp
* 49085 49891: contig of 807 bp in length
* 49892 49991: gap of 100 bp
* 49992 50801: contig of 810 bp in length
* 50802 50901: gap of 100 bp
* 50902 51691: contig of 790 bp in length
* 51692 51791: gap of 100 bp
* 51792 52610: contig of 819 bp in length
* 52611 52710: gap of 100 bp
* 52711 53528: contig of 818 bp in length
* 53529 53628: gap of 100 bp
* 53629 54440: contig of 812 bp in length
* 54441 54540: gap of 100 bp
* 54541 55346: contig of 806 bp in length
* 55347 55446: gap of 100 bp
* 55447 56243: contig of 797 bp in length
* 56244 56343: gap of 100 bp
* 56344 57148: contig of 805 bp in length
* 57149 57248: gap of 100 bp
* 57249 58063: contig of 815 bp in length
* 58064 58163: gap of 100 bp
* 58164 58983: contig of 820 bp in length
* 58984 59083: gap of 100 bp
* 59084 59899: contig of 816 bp in length
* 59900 59999: gap of 100 bp
* 60000 60812: contig of 813 bp in length
* 60813 60912: gap of 100 bp
* 60913 61734: contig of 822 bp in length
* 61735 61834: gap of 100 bp
* 61835 62644: contig of 810 bp in length
* 62645 62744: gap of 100 bp

Query Match 1.3%; Score 40; DB 73; Length 80861;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tcctatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 54838 TCCTACANAAAAAAAAAAAAAAAAAAAAAAAAAAAA 54877

RESULT 20
AC026450/c
LOCUS AC026450 95236 bp DNA HTG 22-MAR-2000
DEFINITION Homo sapiens chromosome 5 clone CTD-2381N15, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
ACCESSION AC026450
VERSION AC026450.1 GI:7279615
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 95236)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 95236)
DOE Joint Genome Institute.
Direct Submission
Submitted (22-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----Summary Statistics
Consensus quality: 83275 bases at least Q40
Consensus quality: 89876 bases at least Q30
Consensus quality: 91237 bases at least Q20
Estimated insert size: 95236; sum-of-contigs estimation
Estimated insert size: 83310; agarose-fp estimation
Quality coverage: 4.77x in Q20 bases; agarose-fp estimation
Quality coverage: 4.17x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1090: contig of 1090 bp in length
* gap of unknown length
* 1091 2567: contig of 1477 bp in length
* gap of unknown length
* 2568 3633: contig of 1066 bp in length
* gap of unknown length
* 3634 5602: contig of 1969 bp in length
* gap of unknown length
* 5603 7519: contig of 1917 bp in length
* gap of unknown length
* 7520 11811: contig of 4292 bp in length
* gap of unknown length
* 11812 17036: contig of 5225 bp in length
* gap of unknown length
* 17037 21227: contig of 4191 bp in length
* gap of unknown length
* 21228 27367: contig of 6140 bp in length
* gap of unknown length
* 27368 35548: contig of 8181 bp in length
* gap of unknown length
* 35549 44349: contig of 8801 bp in length
* gap of unknown length
* 44350 52374: contig of 8025 bp in length
* gap of unknown length
* 52375 63936: contig of 11562 bp in length
* gap of unknown length
* 63937 81454: contig of 17518 bp in length
* gap of unknown length
* 81455 95236: contig of 13782 bp in length.

FEATURES
source

1. .95236
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2381N15"
BASE COUNT 27883 a 19363 c 19396 g 28584 t 10 others
ORIGIN

Query Match 1.3%; Score 40; DB 67; Length 95236;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3078 tcctatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||
Db 2052 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2013

RESULT 21

HSDJ900E8/c

LOCUS

DEFINITION

Human DNA sequence from clone RP5-900E8 on chromosome Xq25-27.1.
Contains the first coding exon of the GPC4 gene for glypican 4,
STSS, GSSs and a putative CpG island, complete sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 101259)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (21-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Apr 9, 2000 this sequence version replaced gi:6015559.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submissions
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
<http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence
was generated from part of bacterial clone contigs of human
Chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/ChrX>
RP5-900E8 is from the library RPCI-5 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see <http://bacpac.med.buffalo.edu/>
VECTOR: pCPAC2

IMPORTANT: This sequence is not the entire insert of clone
RP5-900E8 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP5-900E8 is at 101259 in this
sequence. The true right end of clone RP6-198C21 is at 100 in this
sequence.

FEATURES

source

1..101259
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q25-27.1"
/clone="RP5-900E8"
/clone.lib="RPCI-5"
repeat_region 1..318
/note="MLT1A1 repeat: matches 21..359 of consensus"
repeat_region 395..449
/note="L2 repeat: matches 2634..2689 of consensus"
repeat_region 487..780
/note="AluX repeat: matches 1..295 of consensus"
repeat_region 1050..1352
/note="AluSg repeat: matches 1..302 of consensus"
repeat_region 1363..1652
/note="AluSx repeat: matches 1..299 of consensus"

repeat_region 1889..2184
/note="AluSp repeat: matches 3..304 of consensus"
repeat_region 3252..3558
/note="AluSg repeat: matches 1..307 of consensus"
repeat_region 3577..3872
/note="AluSx repeat: matches 1..294 of consensus"
repeat_region 3922..4005
/note="MER5B repeat: matches 34..117 of consensus"
repeat_region 5247..5532
/note="AluSp repeat: matches 7..299 of consensus"
repeat_region 5666..5956
/note="AluSx repeat: matches 1..300 of consensus"
repeat_region 5957..6134
/note="AluSg/x repeat: matches 132..312 of consensus"
repeat_region 7063..7274
/note="MIR repeat: matches 30..246 of consensus"
repeat_region 8360..8544
/note="L2 repeat: matches 2563..2749 of consensus"
misc_feature complement(9534..9913)
/note="match: GSS: Em:AQ074296"
misc_feature 10324..13040
/note="CpG island"
/evidence=not_experimental
complement(<11590..12194)
/gene="GPC4"
/evidence=not_experimental
/product="dJ900E8.1 (glypican 4)"
complement(11590..12194)
/gene="GPC4"
complement(<11590..11749)
/gene="GPC4"
/codon_start=1
/evidence=not_experimental
/product="dJ900E8.1 (glypican 4)"
/protein_id="CAB8664.1"
/db_xref="GI:7529585"
/translation="MARFGLPALCTLAVLSAALLAELKSKSCVRRLLYVSKGFNK
NDAPLHEIN"
repeat_region 14593..15002
/note="L1MB5 repeat: matches 5767..6173 of consensus"
repeat_region 15310..15354
/note="L2 repeat: matches 2696..2740 of consensus"
repeat_region 15674..15950
/note="AluJb repeat: matches 3..295 of consensus"
repeat_region 15955..16012
/note="29 copies 2 mer gg 69% conserved"
repeat_region 16074..16139
/note="L2 repeat: matches 2086..2150 of consensus"
repeat_region 16166..16283
/note="AluJb repeat: matches 4..124 of consensus"
repeat_region 16302..16505
/note="AluJb repeat: matches 87..310 of consensus"
repeat_region 16632..16768
/note="L2 repeat: matches 2405..2529 of consensus"
repeat_region 16884..17188
/note="AluSx repeat: matches 1..311 of consensus"
repeat_region 17582..18301
/note="MER5B repeat: matches 1..617 of consensus"
repeat_region 18637..18818
/note="L2 repeat: matches 1995..2181 of consensus"
repeat_region 19468..19774
/note="AluJb repeat: matches 1..311 of consensus"
repeat_region 19812..19916
/note="MIR repeat: matches 29..136 of consensus"
repeat_region 20124..20221
/note="L2 repeat: matches 2622..2709 of consensus"
repeat_region 20480..20552
/note="MIR repeat: matches 92..169 of consensus"
repeat_region 20993..21159
/note="AluSg/x repeat: matches 134..301 of consensus"
repeat_region 21160..21465
/note="AluSx repeat: matches 1..305 of consensus"
repeat_region 22657..22795

```
repeat_region /note="AluSg repeat: matches 1. .139 of consensus"
22822. .23144
repeat_region /note="AluJb repeat: matches 1. .308 of consensus"
23557. 23828
repeat_region /note="AluSg1 repeat: matches 5. .301 of consensus"
24093. .24399
repeat_region /note="AluJ repeat: matches 1. .307 of consensus"
24403. .24662
repeat_region /note="MIR repeat: matches 12. .262 of consensus"
24690. 24770
repeat_region /note="LTR15 repeat: matches 172. .248 of consensus"
24790. .25093
repeat_region /note="AluSg repeat: matches 1. .303 of consensus"
25129. .25628
repeat_region /note="L1MB5 repeat: matches 5646. .6143 of consensus"
25637. 25938
repeat_region /note="AluSx repeat: matches 5. .306 of consensus"
25967. .26249
repeat_region /note="AluJo repeat: matches 1. .282 of consensus"
26279. .26567
repeat_region /note="AluJo repeat: matches 7. .281 of consensus"
Complement(26853. .27267)
misc_feature /note="match: GSS: Em:AQ430698"
Complement(26889. .27291)
misc_feature /note="match: GSS: Em:AQ496873"
28401. .28584
repeat_region /note="MER30 repeat: matches 1. .197 of consensus"
29736. .30259
misc_feature /note="match: GSS: Em:B99950"
29945. .30071
repeat_region /note="MIR repeat: matches 34. .165 of consensus"
30073. .30336
repeat_region /note="AluSg repeat: matches 1. .281 of consensus"
30883. .30941
repeat_region /note="AluJb repeat: matches 86. .139 of consensus"
30942. .31601
repeat_region /note="MER51B repeat: matches 1. .617 of consensus"
31602. .31737
repeat_region /note="AluJb repeat: matches 139. .297 of consensus"
31877. .32185
repeat_region /note="AluJo repeat: matches 1. .299 of consensus"
32660. .32766
repeat_region /note="L2 repeat: matches 2634. .2744 of consensus"
33288. .33581
repeat_region /note="AluSx repeat: matches 4. .296 of consensus"
34977. .35285
repeat_region /note="AluJb repeat: matches 7. .310 of consensus"
35360. .35429
repeat_region /note="MIR repeat: matches 93. .162 of consensus"
35750. .35837
repeat_region /note="44 copies 2 mer tt 65% conserved"
35863. .36172
repeat_region /note="AluJo repeat: matches 1. .312 of consensus"
36206. .36502
repeat_region /note="AluSc repeat: matches 1. .308 of consensus"
36921. .37101
repeat_region /note="MIR repeat: matches 76. .255 of consensus"
38901. .39098
repeat_region /note="MER3 repeat: matches 1. .204 of consensus"
39140. .39430
Query Match 1.3%; Score 40; DB 11; Length 101259;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 58491 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 58452
```

```
RESULT 22
AC020928
LOCUS AC020928 103804 bp DNA HTG 06-APR-2000
```

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Homo sapiens chromosome 19 clone CTD-2162K18, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
AC020928
AC020928.3 GI:7458788
HTG: HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 103804)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 19
Unpublished
2 (bases 1 to 103804)
DOE Joint Genome Institute.
Direct Submission
Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 6, 2000 this sequence version replaced gi:7021678.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----Summary Statistics
Consensus quality: 96332 bases at least Q40
Consensus quality: 101562 bases at least Q30
Consensus quality: 102259 bases at least Q20
Estimated insert size: 103804; sum-of-contigs estimation
Estimated insert size: 111350; agarose-fp estimation
Quality coverage: 4.80x in Q20 bases; agarose-fp estimation
Quality coverage: 5.15x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 2062: contig of 2062 bp in length
* gap of unknown length
* 6764: contig of 4702 bp in length
* gap of unknown length
* 6765 12879: contig of 6115 bp in length
* gap of unknown length
* 12880 19068: contig of 6189 bp in length
* gap of unknown length
* 19069 23463: contig of 4395 bp in length
* gap of unknown length
* 23464 35319: contig of 11856 bp in length
* gap of unknown length
* 35320 47360: contig of 12041 bp in length
* gap of unknown length
* 47361 61652: contig of 14292 bp in length
* gap of unknown length
* 61653 81957: contig of 20305 bp in length
* gap of unknown length
* 81958 103804: contig of 21847 bp in length.
* Location/Qualifiers
1. .103804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2162K18"
BASE COUNT 30040 a 22650 c 22148 g 28963 t 3 others
ORIGIN

FEATURES
source

Query Match 1.3%; Score 40; DB 73; Length 103804;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 92154 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 92193

RESULT 23

AC007880 AC007880 122223 bp DNA PRI 21-DEC-1999
LOCUS Homo sapiens BAC clone RP11-323F11 from 2, complete sequence.
DEFINITION AC007880
ACCESSION AC007880
VERSION AC007880.2 GI:5931460
KEYWORDS HTG.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 122223)
Sulston, J.E. and Waterston, R.

AUTHORS

Toward a complete human genome sequence
99063792

MEDLINE

2 (bases 1 to 122223)
Du, H., Maupin, R. A. C. H. E. L., Yoakum, M. and Nguyen, C.

TITLE

The sequence of Homo sapiens BAC clone RP11-323F11

JOURNAL

Unpublished

REFERENCE

3 (bases 1 to 122223)
Waterston, R. H.

AUTHORS

Direct Submission

TITLE

Submitted (19-JUN-1999) Genome Sequencing Center, Washington

JOURNAL

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

4 (bases 1 to 122223)
Waterston, R. H.

AUTHORS

Direct Submission

TITLE

Submitted (28-SEP-1999) Genome Sequencing Center, Washington

JOURNAL

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

5 (bases 1 to 122223)
Waterston, R.

AUTHORS

Direct Submission

TITLE

Submitted (21-DEC-1999) Department of Genetics, Washington

JOURNAL

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE

On Sep 28, 1999 this sequence version replaced gi:5103899.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0323F11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis

MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-359K10, 200 bp overlap; the clone sequenced to the right is RP11-451M22. Actual start of this clone is at base position 147528 of RP11-359K10.

FEATURES

source

1. .122223
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-323F11"
/clone_lib="RPCI-11"
694. .804
/rpt_family="MER1_type"
1020. .1322
/rpt_family="Alu"
1367. .1661
/rpt_family="Alu"
1792. .1960
/rpt_family="Retroviral"
1957. .2161
/rpt_family="Retroviral"
4727. .4751
/rpt_family="(TTA)n"
4752. .10803
/rpt_family="L1"
13379. .13905
/rpt_family="Retroviral"
13902. .13962
/rpt_family="Harlequin"
13959. .14032
/rpt_family="Harlequin"
14033. .14196
/rpt_family="Retroviral"
14200. .14318
/rpt_family="Retroviral"
14305. .14415
/rpt_family="Retroviral"
15049. .15114
/rpt_family="MER1_type"
15105. .15485
/note="match to EST AI809723 (NID:G5396289) wh77a01.x1"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region	17318. .17623	/rpt_family="Alu"
repeat_region	17624. .17693	/rpt_family="MaLR"
repeat_region	17694. .17806	/rpt_family="L1"
repeat_region	17807. .17829	/rpt_family="(TTAA)n"
repeat_region	17830. .18105	/rpt_family="Alu"
repeat_region	18106. .18159	/rpt_family="L1"
repeat_region	18227. .18525	/rpt_family="Alu"
misc_feature	19214. .19349	/note="match to EST T81027 (NID:g703912) yd25c02.st"
repeat_region	19310. .19420	/rpt_family="L1"
repeat_region	19421. .19704	/rpt_family="Alu"
repeat_region	19705. .19747	/rpt_family="L1"
repeat_region	19770. .20037	/rpt_family="L2"
repeat_region	20178. .20286	/rpt_family="L2"
repeat_region	20304. .20449	/rpt_family="MER2_type"
repeat_region	20721. .21260	/rpt_family="Retroviral"
repeat_region	21790. .21899	/rpt_family="MIR"
repeat_region	22010. .22404	/rpt_family="L2"
repeat_region	23178. .23398	/rpt_family="L1"
repeat_region	23399. .23826	/rpt_family="LTR34"
repeat_region	23827. .24134	/rpt_family="Alu"
repeat_region	24135. .24243	/rpt_family="L1"
repeat_region	24244. .24471	/rpt_family="L1"
repeat_region	24472. .24764	/rpt_family="Alu"
repeat_region	24765. .24888	/rpt_family="L1"
repeat_region	24907. .25029	/rpt_family="Alu"
repeat_region	25711. .25787	/rpt_family="MER1_type"
repeat_region	26651. .26675	/rpt_family="AT-rich"
repeat_region	26866. .26926	/rpt_family="Alu"
repeat_region	26944. .26985	/rpt_family="(GAAA)n"
repeat_region	26985. .27031	/rpt_family="(GGGA)n"
repeat_region	27824. .27945	

```
Query Match      1.3%; Score 40; DB 39; Length 122223;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaa 3117
|||
Db 70964 TCCATCAAAAAAAAAAAAAAAAAAAAAA 71003

Db 70964 TCCATCAAAAAAAAAAAAAAAAAAAAAA 71003

RESULT 24
AC008932/C

```

LOCUS          AC008932   128758 bp   DNA               05-APR-2000
DEFINITION     Homo sapiens chromosome 5 clone CTD-2296H2, WORKING DRAFT SEQUENCE,
                7 unordered pieces.
ACCESSION      AC008932
VERSION        AC008932.3   GI:7417522
KEYWORDS       HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 128758)
AUTHORS       DOE Joint Genome Institute.
TITLE         Sequencing of Human Chromosome 5
JOURNAL       Unpublished
REFERENCE      2 (bases 1 to 128758)
AUTHORS       DOE Joint Genome Institute.
TITLE         Direct Submission
JOURNAL       Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
                Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT       On Apr 5, 2000 this sequence version replaced gi:7025715.
                -----Genome Center
                Center: Joint Genome Institute
                Center Code: JGI
                Web site: http://www.jgi.doe.gov
                -----Summary Statistics
                Consensus quality: 122924 bases at least Q40
                Consensus quality: 126992 bases at least Q30
                Consensus quality: 127709 bases at least Q20
                Estimated insert size: 128758; Sum-of-contigs estimation
                Estimated insert size: 131000; pulse field gel estimation
                Quality coverage: 5.00x in Q20 bases; pulse field gel estimation
                Quality coverage: 5.09x in Q20 bases; sum-of-contigs estimation
                -----
                * NOTE: This is a 'working draft' sequence. It currently
                * consists of 7 contigs. The true order of the pieces
                * is not known and their order in this sequence record is
                * arbitrary. Gaps between the contigs are represented as
                * runs of N, but the exact sizes of the gaps are unknown.
                * This record will be updated with the finished sequence
                * as soon as it is available and the accession number will
                * be preserved.
                *
                * 1 4080: contig of 4080 bp in length
                * gap of unknown length
                * 4081 9811: contig of 5731 bp in length
                * gap of unknown length
                * 9812 17645: contig of 7834 bp in length
                * gap of unknown length
                * 17646 26939: contig of 9294 bp in length
                * gap of unknown length
                * 26940 47546: contig of 20607 bp in length
                * gap of unknown length
                * 47547 69345: contig of 21799 bp in length
                * gap of unknown length
                * 69346 128758: contig of 59413 bp in length.

```

BASE COUNT	41820 a	22954 c	23096 g	40881 t	7 others
ORIGIN					

```
Query Match      1.3%; Score 40; DB 72; Length 128758;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	3078	tccatcaaaaaaaaaaaaaaaaaaaaaaa	3117
Dβ	71426	TCCATCAAAAAAAAAAAAAAAAAAAAAA	7138

Db 71426 TCCATCAAAAAAAAAAAAAAAAAAAAAA 71387

RESULT 25

HS198C21/c

LOCUS

HS198C21 141762 bp DNA PRI 23-NOV-1999
Human DNA sequence from clone 198C21 on chromosome Xq26.1-26.3
Contains GPC4 (glypican 4), ESTs, STSS and GSS, complete sequence.

ACCESSION

AL034400

VERSION

AL034400.2 GI:4455461

KEYWORDS

HTG; GPC4.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 141762)

REFERENCE

Heath, P.

AUTHORS

Direct Submission

TITLE

Submitted (16-MAR-1999) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Mar 21, 1999 this sequence version replaced gi:3980345.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

IMPORTANT: This sequence is not the entire insert of clone 198C21.

It may be shorter because we only sequence overlapping sections

once, or longer because we arrange for a small overlap between

neighbouring submissions.

The true right end of clone dj358H7 is at 100 in this sequence.

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

This sequence was generated from part of bacterial clone contigs of

human chromosome X, constructed by the Sanger Centre Chromosome X

Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/ChrX

198C21 is from the library RPC16 constructed at the Roswell Park

Cancer Institute by the group of Pieter de Jong. For further

details see http://bacpac.med.buffalo.edu/VECTOR: ppAC4.

Location/Qualifiers

1..141762

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="q26.1-26.3"

/clone="RP6-198C21"

/clone_lib="RPC1-6"

1..397

/note="MLR1D repeat: matches 100. .500 of consensus"

757. .810

/note="27 copies 2 mer tc 85% conserved"

767. .810

/note="11 copies 4 mer tctc 93% conserved"

813. .1075

/note="AluJb repeat: matches 1. .274 of consensus"

1166. .1204

/note="THE1B repeat: matches 326. .364 of consensus"

1205. .1504

/note="AluSg repeat: matches 1. .301 of consensus"

1505. .1830

/note="THE1B repeat: matches 1. .326 of consensus"

1974. .2325

/note="match: 267843 STS containing (CA) repeat"

2115. .2143

/note="14 copies 2 mer ca 100% conserved; differs from

267843"

2263. .2496

/note="MIR repeat: matches 20. .262 of consensus"

2498. .2782

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/note="AluSg repeat: matches 1. .293 of consensus"
2808. .3085
/note="AluJo repeat: matches 1. .278 of consensus"
3091. .3126
/note="9 copies 4 mer aata 83% conserved"
3303. .3435
/note="FLAM_A repeat: matches 1. .133 of consensus"
4563. .4594
/note="16 copies 2 mer aa 91% conserved"
4734. .4771
/note="19 copies 2 mer tt 79% conserved"
4851. .4963
/note="SVA repeat: matches 764. .870 of consensus"
4972. .5123
/note="AluX repeat: matches 126. .287 of consensus"
5124. .5419
/note="AluY repeat: matches 1. .296 of consensus"
5420. .5548
/note="AluX repeat: matches 1. .126 of consensus"
5716. .5767
/note="13 copies 4 mer cctg 81% conserved"
5928. .6249
/note="MER33 repeat: matches 3. .324 of consensus"
8348. .8438
/note="L2 repeat: matches 2598. .2700 of consensus"
8620. .8744
/note="L2 repeat: matches 2588. .2710 of consensus"
10074. .10378
/note="AluSg repeat: matches 1. .299 of consensus"
10383. .10517
/note="AluJb repeat: matches 1. .142 of consensus"
10570. .10796
/note="MER93 repeat: matches 161. .397 of consensus"
complement(11387. .11797)
/note="match: GSS AQ311874 clone R-10115"
complement(11911. .12074)
/note="match: STS I42703"
12692. .12804
/note="L1PB3 repeat: matches 6035. .6147 of consensus"
14446. .14489
/note="11 copies 4 mer gaat 77% conserved"
14629. .14920
/note="AluJo repeat: matches 1. .291 of consensus"
14971. .15016
/note="L2 repeat: matches 2651. .2696 of consensus"
15708. .16063
/note="THE1B repeat: matches 1. .362 of consensus"
16472. .16571
/note="MERSA repeat: matches 9. .112 of consensus"
16648. .16836
/note="MIR repeat: matches 47. .237 of consensus"
16866. .17136
/note="AluJb repeat: matches 24. .301 of consensus"
17137. .23554
/note="L1PA8 repeat: matches 1. .6159 of consensus"
24563. .24592
/note="10 copies 3 mer tga 97% conserved"
24595. .24894
/note="AluJo repeat: matches 1. .286 of consensus"
24898. .25029
/note="AluJo/FLAM repeat: matches 1. .133 of consensus"
25356. .25652
/note="AluX repeat: matches 1. .297 of consensus"
26706. .26990
/note="AluJb repeat: matches 11. .302 of consensus"
27545. .27799
/note="MLT1J repeat: matches 258. .515 of consensus"
28044. .28342
/note="AluSg repeat: matches 1. .299 of consensus"
28625. .28770
/note="MERSA repeat: matches 1. .154 of consensus"
28879. .29178
/note="AluX repeat: matches 1. .296 of consensus"

```
repeat_region 29218..29347
/note="L2 repeat: matches 2574. .2703 of consensus"
repeat_region 29474..29546
/note="MSTD repeat: matches 327. .394 of consensus"
repeat_region 29547..29859
/note="AluJb repeat: matches 1. .306 of consensus"
repeat_region 29860..30156
/note="MSTD repeat: matches 1. .327 of consensus"
repeat_region 33163..33464
/note="AluSx repeat: matches 1. .302 of consensus"
repeat_region 34349..34666
/note="AluSx repeat: matches 1. .312 of consensus"
misc_feature complement(<34667..>34935)
/note="match: GSS B43565"
misc_feature complement(<34823..>35148)
/note="match: GSS AQ169413"
repeat_region 35018..35130
/note="MIR repeat: matches 33. .143 of consensus"
repeat_region 35216..35521
/note="AluSg1 repeat: matches 1. .303 of consensus"
repeat_region 36157..36204
/note="12 copies 4 mer acac 94% conserved"
repeat_region 36157..36202
/note="23 copies 2 mer ac 96% conserved"
misc_feature 36453..36901
/note="match: GSS AQ265693 clone 2510C6"
misc_feature 36453..36748
/note="match: GSS AQ265745 clone 2510C22"
misc_feature 36453..36751
/note="match: GSS AQ263897 clone 2510D5"
repeat_region 36765..37053
/note="AluJo repeat: matches 1. .289 of consensus"
repeat_region 37520..37831
/note="AluSx repeat: matches 5. .312 of consensus"
prim_transcript 38233..38731
/note="match: 3' EST AI417996 clone IMAGE:2114220"
repeat_region 38378..38401
/note="12 copies 2 mer ca 96% conserved"
repeat_region 38429..38481
/note="MADE1 repeat: matches 1. .53 of consensus"
prim_transcript 38998..39479
/note="match: 3' EST AI368019 clone IMAGE:1934935"
prim_transcript 38998..39488
/note="match: 3' EST AI276057 clone IMAGE:1877575"
mRNA complement(join(<39482..41515,41612..41787,43171..43307,
44218..44364,44470..44600,49704..49869,62591..62982,
77629..>77787))
/note="match: AF030186; match: ESTs AA046087 AA046130
AA478468 AB93150 AA933954 AI071251 AI131068 AI141502
AI208484 AI385680 C82918 C83774 D19430 D25838 W12478
AA451958 C82589 C83445 C82710 C83566 W79980"
/evidence=not_experimental
repeat_region 40979..41026
/note="12 copies 4 mer acaa 75% conserved"
repeat_region 40979..41046
/note="34 copies 2 mer aa 66% conserved"
```

Query Match 1.3%; Score 40; DB 10; Length 141762;
Best local Similarity 100.0%; Pred. No. 1.le-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 82696 TCCATCANNAAAAAAAAAAAAAAAAAAAAAAAAAAAA 82657

RESULT 26
AC034243/c AC034243 DNA HTG 05-APR-2000
LOCUS Homo sapiens chromosome 5 clone RPI-204011, LOW-PASS SEQUENCE
DEFINITION SAMPLING.
AC034243
ACCESSION AC034243
VERSION AC034243.1 GI:7417720
```

HTG; HTGS_PHASE0.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DOE Joint Genome Institute.
1 (bases 1 to 166947)
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 166947)
DOE Joint Genome Institute.
Direct Submission
Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

* NOTE: This record contains 191 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 572: contig of 572 bp in length
* gap of unknown length
* 573 1166: contig of 594 bp in length
* gap of unknown length
* 1167 1829: contig of 663 bp in length
* gap of unknown length
* 1830 2234: contig of 405 bp in length
* gap of unknown length
* 2235 2359: contig of 125 bp in length
* gap of unknown length
* 2360 2960: contig of 601 bp in length
* gap of unknown length
* 2961 3545: contig of 585 bp in length
* gap of unknown length
* 3546 3915: contig of 370 bp in length
* gap of unknown length
* 3916 4482: contig of 567 bp in length
* gap of unknown length
* 4483 5243: contig of 761 bp in length
* gap of unknown length
* 5244 5861: contig of 618 bp in length
* gap of unknown length
* 5862 6572: contig of 711 bp in length
* gap of unknown length
* 6573 7297: contig of 725 bp in length
* gap of unknown length
* 7298 7450: contig of 153 bp in length
* gap of unknown length
* 7451 8131: contig of 681 bp in length
* gap of unknown length
* 8132 8552: contig of 421 bp in length
* gap of unknown length
* 8553 9149: contig of 597 bp in length
* gap of unknown length
* 9150 9943: contig of 794 bp in length
* gap of unknown length
* 9944 10502: contig of 559 bp in length
* gap of unknown length
* 10503 10878: contig of 376 bp in length
* gap of unknown length
* 10879 11524: contig of 646 bp in length
* gap of unknown length
* 11525 11885: contig of 361 bp in length


```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 167627)
Waterston,R.H.
Direct Submission
Submitted (17-NOV-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Apr 21, 2000 this sequence version replaced gi:6524380.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0133L20
----- Summary Statistics -----
Sequencing vector: M13; 84%
Sequencing vector: plasmid; 16%
Chemistry: Dye-primer ET; 84% of reads
Chemistry: Dye-terminator Big Dye; 16% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165708 bases at least Q40
Consensus quality: 166084 bases at least Q30
Consensus quality: 166501 bases at least Q20
Insert size: 163000; agarose-fp
Insert size: 167527; sum-of-contigs
Quality coverage: 9.28 in Q20 bases; agarose-fp
Quality coverage: 9.07 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 11767: contig of 11767 bp in length
* 11768 11867: gap of unknown length
* 11868 167627: contig of 155760 bp in length.

FEATURES
Source
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/clone="RP11-133L20"

BASE COUNT 46371 a 38471 c 38705 g 43980 t 100 others
ORIGIN

Query Match 1.3%; Score 40; DB 78; Length 167627;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 65760 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 65799

RESULT 28
AC007465 170125 bp DNA HTG 05-JUN-1999
LOCUS Homo sapiens clone NH0575121, *** SEQUENCING IN PROGRESS ***, 1
DEFINITION unordered pieces.
ACCESSION AC007465
VERSION AC007465.2 GI:5001471
KEYWORDS HTG: HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 170125)

Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 170125)
Waterston,R.H.
Direct Submission
Submitted (04-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 5, 1999 this sequence version replaced gi:4734032.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 170125: contig of 170125 bp in length.

FEATURES
Source
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NH0575121"

BASE COUNT 53393 a 32219 c 31961 g 52552 t
ORIGIN

Query Match 1.3%; Score 40; DB 41; Length 170125;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 113571 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 113610

RESULT 29
AL139412 170273 bp DNA HTG 19-FEB-2000
LOCUS Homo sapiens chromosome 1 clone RP11-98G7 map q21.2-22, ***
DEFINITION SEQUENCING IN PROGRESS ***, 11 unordered pieces.
ACCESSION AL139412
VERSION AL139412.2 GI:7009514
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 170273)

Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 170125)
Waterston,R.H.
Direct Submission
Submitted (19-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Feb 20, 2000 this sequence version replaced gi:6996233.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in
progress and the release of this data is based on the understanding
that the sequence may change as work continues. The sequence may
be contaminated with foreign sequence from E.coli, yeast, vector,
phage etc. Order of segments is not known; 800 n's separate
segments. Contig_ID: 00007 Length: 2124bp
Contig_ID: 00019 Length: 21300bp
Contig_ID: 00130 Length: 3263bp
Contig_ID: 00270 Length: 13349bp
Contig_ID: 00664 Length: 1257bp
Contig_ID: 00693 Length: 13321bp
Contig_ID: 00699 Length: 12395bp
Contig_ID: 00740 Length: 18584bp
Contig_ID: 00912 Length: 67931bp
Contig_ID: 01146 Length: 5013bp
Contig_ID: 01496 Length: 3736bp.
```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 2124: contig of 2124 bp in length
* 2125 2924: gap of 800 bp
* 2925 24224: contig of 21300 bp in length
* 24225 25024: gap of 800 bp
* 25025 28287: contig of 3263 bp in length
* 28288 29087: gap of 800 bp
* 29088 42436: contig of 13349 bp in length
* 42437 43236: gap of 800 bp
* 43237 44493: contig of 1257 bp in length
* 44494 45293: gap of 800 bp
* 45294 58614: contig of 13321 bp in length
* 58615 59414: gap of 800 bp
* 59415 71809: contig of 12395 bp in length
* 71810 72609: gap of 800 bp
* 72610 91193: contig of 18584 bp in length
* 91194 91993: gap of 800 bp
* 91994 159924: contig of 67931 bp in length
* 159925 160724: gap of 800 bp
* 160725 165737: contig of 5013 bp in length
* 165738 166537: gap of 800 bp
* 166538 170273: contig of 3736 bp in length.

FEATURES
source

1. .170273
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="q21.2-22"
/clone="RP11-98G7"
/clone_lib="RPC1-11.1"
BASE COUNT 40295 a 41235 c 40421 g 40317 t 8005 others
ORIGIN

Query Match 1.3%; Score 40; DB 32; Length 170273;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 65755 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 65794

RESULT 30
AC012312

LOCUS AC012312 170948 bp DNA HTG 05-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTC-558019, WORKING DRAFT SEQUENCE,
24 unordered pieces.

ACCESSION AC012312
VERSION AC012312.3 GI:7417591
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 170948)
DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170948)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission

Submitted (23-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 5, 2000 this sequence version replaced gi:6693123.
-----Genome Center

FEATURES
source

1. .170948
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-558019"

Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----Summary Statistics
Consensus quality: 136918 bases at least Q40
Consensus quality: 157435 bases at least Q30
Consensus quality: 162969 bases at least Q20
Estimated insert size: 170948; sum-of-contigs estimation
Estimated insert size: 169000; pulse field gel estimation
Quality coverage: 3.92x in Q20 bases; pulse field gel estimation
Quality coverage: 3.88x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 2634: contig of 2634 bp in length
* 2635 4663: gap of unknown length
* 4664 6764: contig of 2029 bp in length
* 6765 8848: gap of unknown length
* 8849 11282: contig of 2101 bp in length
* 11283 13644: gap of unknown length
* 13645 16272: contig of 2084 bp in length
* 16273 18406: gap of unknown length
* 18407 21475: contig of 2434 bp in length
* 21476 23845: gap of unknown length
* 23846 26839: contig of 2134 bp in length
* 26840 32350: gap of unknown length
* 32351 37037: contig of 2994 bp in length
* 37038 42273: contig of 5511 bp in length
* 42274 47489: gap of unknown length
* 47490 53940: contig of 4687 bp in length
* 53941 60016: gap of unknown length
* 60017 68037: contig of 5236 bp in length
* 68038 82025: gap of unknown length
* 82026 92693: contig of 6076 bp in length
* 92694 106625: contig of 8021 bp in length
* 106626 122562: gap of unknown length
* 122563 144230: contig of 10668 bp in length
* 144231 170948: contig of 13932 bp in length
* 170949 21668: gap of unknown length
* 21669 26718: contig of 15937 bp in length
* 26719 28000: gap of unknown length
* 28001 30000: contig of 21668 bp in length
* 30001 32000: gap of unknown length
* 32002 34000: contig of 15937 bp in length
* 34003 36000: gap of unknown length
* 36004 38000: contig of 13988 bp in length
* 38005 40000: gap of unknown length
* 40006 42000: contig of 10668 bp in length
* 42007 44000: gap of unknown length
* 44008 46000: contig of 13932 bp in length
* 46009 48000: gap of unknown length
* 48010 50000: contig of 15937 bp in length
* 50011 52000: gap of unknown length
* 52012 54000: contig of 13988 bp in length
* 54013 56000: gap of unknown length
* 56014 58000: contig of 10668 bp in length
* 58015 60000: gap of unknown length
* 60016 62000: contig of 13932 bp in length
* 62017 64000: gap of unknown length
* 64018 66000: contig of 15937 bp in length
* 66019 68000: gap of unknown length
* 68020 70000: contig of 13988 bp in length
* 70021 72000: gap of unknown length
* 72022 74000: contig of 10668 bp in length
* 74023 76000: gap of unknown length
* 76024 78000: contig of 13932 bp in length
* 78025 80000: gap of unknown length
* 80026 82000: contig of 15937 bp in length
* 82027 84000: gap of unknown length
* 84028 86000: contig of 13988 bp in length
* 86029 88000: gap of unknown length
* 88030 90000: contig of 10668 bp in length
* 90031 92000: gap of unknown length
* 92032 94000: contig of 13932 bp in length
* 94033 96000: gap of unknown length
* 96034 98000: contig of 15937 bp in length
* 98035 100000: gap of unknown length
* 100036 102000: contig of 13988 bp in length
* 102037 104000: gap of unknown length
* 104038 106000: contig of 10668 bp in length
* 106039 108000: gap of unknown length
* 108040 110000: contig of 13932 bp in length
* 110041 112000: gap of unknown length
* 112042 114000: contig of 15937 bp in length
* 114043 116000: gap of unknown length
* 116044 118000: contig of 13988 bp in length
* 118045 120000: gap of unknown length
* 120046 122000: contig of 10668 bp in length
* 122047 124000: gap of unknown length
* 124048 126000: contig of 13932 bp in length
* 126049 128000: gap of unknown length
* 128050 130000: contig of 15937 bp in length
* 130051 132000: gap of unknown length
* 132052 134000: contig of 13988 bp in length
* 134053 136000: gap of unknown length
* 136054 138000: contig of 10668 bp in length
* 138055 140000: gap of unknown length
* 140056 142000: contig of 13932 bp in length
* 142057 144000: gap of unknown length
* 144058 146000: contig of 15937 bp in length
* 146059 148000: gap of unknown length
* 148060 150000: contig of 13988 bp in length
* 150061 152000: gap of unknown length
* 152062 154000: contig of 10668 bp in length
* 154063 156000: gap of unknown length
* 156064 158000: contig of 13932 bp in length
* 158065 160000: gap of unknown length
* 160066 162000: contig of 15937 bp in length
* 162067 164000: gap of unknown length
* 164068 166000: contig of 13988 bp in length
* 166069 168000: gap of unknown length
* 168070 170000: contig of 10668 bp in length
* 170071 172000: gap of unknown length
* 172072 174000: contig of 13932 bp in length
* 174073 176000: gap of unknown length
* 176074 178000: contig of 15937 bp in length
* 178075 180000: gap of unknown length
* 180076 182000: contig of 13988 bp in length
* 182077 184000: gap of unknown length
* 184078 186000: contig of 10668 bp in length
* 186079 188000: gap of unknown length
* 188080 190000: contig of 13932 bp in length
* 190081 192000: gap of unknown length
* 192082 194000: contig of 15937 bp in length
* 194083 196000: gap of unknown length
* 196084 198000: contig of 13988 bp in length
* 198085 200000: gap of unknown length
* 200086 202000: contig of 10668 bp in length
* 202087 204000: gap of unknown length
* 204088 206000: contig of 13932 bp in length
* 206089 208000: gap of unknown length
* 208090 210000: contig of 15937 bp in length
* 210091 212000: gap of unknown length
* 212092 214000: contig of 13988 bp in length
* 214093 216000: gap of unknown length
* 216094 218000: contig of 10668 bp in length
* 218095 220000: gap of unknown length
* 220096 222000: contig of 13932 bp in length
* 222097 224000: gap of unknown length
* 224098 226000: contig of 15937 bp in length
* 226099 228000: gap of unknown length
* 228100 230000: contig of 13988 bp in length
* 230101 232000: gap of unknown length
* 232102 234000: contig of 10668 bp in length
* 234103 236000: gap of unknown length
* 236104 238000: contig of 13932 bp in length
* 238105 240000: gap of unknown length
* 240106 242000: contig of 15937 bp in length
* 242107 244000: gap of unknown length
* 244108 246000: contig of 13988 bp in length
* 246109 248000: gap of unknown length
* 248110 250000: contig of 10668 bp in length
* 250111 252000: gap of unknown length
* 252112 254000: contig of 13932 bp in length
* 254113 256000: gap of unknown length
* 256114 258000: contig of 15937 bp in length
* 258115 260000: gap of unknown length
* 260116 262000: contig of 13988 bp in length
* 262117 264000: gap of unknown length
* 264118 266000: contig of 10668 bp in length
* 266119 268000: gap of unknown length
* 268120 270000: contig of 13932 bp in length
* 270121 272000: gap of unknown length
* 272122 274000: contig of 15937 bp in length
* 274123 276000: gap of unknown length
* 276124 278000: contig of 13988 bp in length
* 278125 280000: gap of unknown length
* 280126 282000: contig of 10668 bp in length
* 282127 284000: gap of unknown length
* 284128 286000: contig of 13932 bp in length
* 286129 288000: gap of unknown length
* 288130 290000: contig of 15937 bp in length
* 290131 292000: gap of unknown length
* 292132 294000: contig of 13988 bp in length
* 294133 296000: gap of unknown length
* 296134 298000: contig of 10668 bp in length
* 298135 300000: gap of unknown length
* 300136 302000: contig of 13932 bp in length
* 302137 304000: gap of unknown length
* 304138 306000: contig of 15937 bp in length
* 306139 308000: gap of unknown length
* 308140 310000: contig of 13988 bp in length
* 310141 312000: gap of unknown length
* 312142 314000: contig of 10668 bp in length
* 314143 316000: gap of unknown length
* 316144 318000: contig of 13932 bp in length
* 318145 320000: gap of unknown length
* 320146 322000: contig of 15937 bp in length
* 322147 324000: gap of unknown length
* 324148 326000: contig of 13988 bp in length
* 326149 328000: gap of unknown length
* 328150 330000: contig of 10668 bp in length
* 330151 332000: gap of unknown length
* 332152 334000: contig of 13932 bp in length
* 334153 336000: gap of unknown length
* 336154 338000: contig of 15937 bp in length
* 338155 340000: gap of unknown length
* 340156 342000: contig of 13988 bp in length
* 342157 344000: gap of unknown length
* 344158 346000: contig of 10668 bp in length
* 346159 348000: gap of unknown length
* 348160 350000: contig of 13932 bp in length
* 350161 352000: gap of unknown length
* 352162 354000: contig of 15937 bp in length
* 354163 356000: gap of unknown length
* 356164 358000: contig of 13988 bp in length
* 358165 360000: gap of unknown length
* 360166 362000: contig of 10668 bp in length
* 362167 364000: gap of unknown length
* 364168 366000: contig of 13932 bp in length
* 366169 368000: gap of unknown length
* 368170 370000: contig of 15937 bp in length
* 370171 372000: gap of unknown length
* 372172 374000: contig of 13988 bp in length
* 374173 376000: gap of unknown length
* 376174 378000: contig of 10668 bp in length
* 378175 380000: gap of unknown length
* 380176 382000: contig of 13932 bp in length
* 382177 384000: gap of unknown length
* 384178 386000: contig of 15937 bp in length
* 386179 388000: gap of unknown length
* 388180 390000: contig of 13988 bp in length
* 390181 392000: gap of unknown length
* 392182 394000: contig of 10668 bp in length
* 394183 396000: gap of unknown length
* 396184 398000: contig of 13932 bp in length
* 398185 400000: gap of unknown length
* 400186 402000: contig of 15937 bp in length
* 402187 404000: gap of unknown length
* 404188 406000: contig of 13988 bp in length
* 406189 408000: gap of unknown length
* 408190 410000: contig of 10668 bp in length
* 410191 412000: gap of unknown length
* 412192 414000: contig of 13932 bp in length
* 414193 416000: gap of unknown length
* 416194 418000: contig of 15937 bp in length
* 418195 420000: gap of unknown length
* 420196 422000: contig of 13988 bp in length
* 422197 424000: gap of unknown length
* 424198 426000: contig of 10668 bp in length
* 426199 428000: gap of unknown length
* 428200 430000: contig of 13932 bp in length
* 430201 432000: gap of unknown length
* 432202 434000: contig of 15937 bp in length
* 434203 436000: gap of unknown length
* 436204 438000: contig of 13988 bp in length
* 438205 440000: gap of unknown length
* 440206 442000: contig of 10668 bp in length
* 442207 444000: gap of unknown length
* 444208 446000: contig of 13932 bp in length
* 446209 448000: gap of unknown length
* 448210 450000: contig of 15937 bp in length
* 450211 452000: gap of unknown length
* 452212 454000: contig of 13988 bp in length
* 454213 456000: gap of unknown length
* 456214 458000: contig of 10668 bp in length
* 458215 460000: gap of unknown length
* 460216 462000: contig of 13932 bp in length
* 462217 464000: gap of unknown length
* 464218 466000: contig of 15937 bp in length
* 466219 468000: gap of unknown length
* 468220 470000: contig of 13988 bp in length
* 470221 472000: gap of unknown length
* 472222 474000: contig of 10668 bp in length
* 474223 476000: gap of unknown length
* 476224 478000: contig of 13932 bp in length
* 478225 480000: gap of unknown length
* 480226 482000: contig of 15937 bp in length
* 482227 484000: gap of unknown length
* 484228 486000: contig of 13988 bp in length
* 486229 488000: gap of unknown length
* 488230 490000: contig of 10668 bp in length
* 490231 492000: gap of unknown length
* 492232 494000: contig of 13932 bp in length
* 494233 496000: gap of unknown length
* 496234 498000: contig of 15937 bp in length
* 498235 500000: gap of unknown length
* 500236 502000: contig of 13988 bp in length
* 502237 504000: gap of unknown length
* 504238 506000: contig of 10668 bp in length
* 506239 508000: gap of unknown length
* 508240 510000: contig of 13932 bp in length
* 510241 512000: gap of unknown length
* 512242 514000: contig of 15937 bp in length
* 514243 516000: gap of unknown length
* 516244 518000: contig of 13988 bp in length
* 518245 520000: gap of unknown length
* 520246 522000: contig of 10668 bp in length
* 522247 524000: gap of unknown length
* 524248 526000: contig of 13932 bp in length
* 526249 528000: gap of unknown length
* 528250 530000: contig of 15937 bp in length
* 530251 532000: gap of unknown length
* 532252 534000: contig of 13988 bp in length
* 534253 536000: gap of unknown length
* 536254 538000: contig of 10668 bp in length
* 538255 540000: gap of unknown length
* 540256 542000: contig of 13932 bp in length
* 542257 544000: gap of unknown length
* 544258 546000: contig of 15937 bp in length
* 546259 548000: gap of unknown length
* 548260 550000: contig of 13988 bp in length
* 550261 552000: gap of unknown length
* 552262 554000: contig of 10668 bp in length
* 554263 556000: gap of unknown length
* 556264 558000: contig of 13932 bp in length
* 558265 560000: gap of unknown length
* 560266 562000: contig of 15937 bp in length
* 562267 564000: gap of unknown length
* 564268 566000: contig of 13988 bp in length
* 566269 568000: gap of unknown length
* 568270 570000: contig of 10668 bp in length
* 570271 572000: gap of unknown length
* 572272 574000: contig of 13932 bp in length
* 574273 576000: gap of unknown length
* 576274 578000: contig of 15937 bp in length
* 578275 580000: gap of unknown length
* 580276 582000: contig of 13988 bp in length
* 582277 584000: gap of unknown length
* 584278 586000: contig of 10668 bp in length
* 586279 588000: gap of unknown length
* 588280 590000: contig of 13932 bp in length
* 590281 592000: gap of unknown length
* 592282 594000: contig of 15937 bp in length
* 594283 596000: gap of unknown length
* 596284 598000: contig of 13988 bp in length
* 598285 600000: gap of unknown length
* 600286 602000: contig of 10668 bp in length
* 602287 604000: gap of unknown length
* 604288 606000: contig of 13932 bp in length
* 606289 608000: gap of unknown length
* 608290 610000: contig of 15937 bp in length
* 610291 612000: gap of unknown length
* 612292 614000: contig of 13988 bp in length
* 614293 616000: gap of unknown length
* 616294 618000: contig of 10668 bp in length
* 618295 620000: gap of unknown length
* 620296 622000: contig of 13932 bp in length
* 622297 624000: gap of unknown length
* 624298 626000: contig of 15937 bp in length
* 626299 628000: gap of unknown length
* 628300 630000: contig of 13988 bp in length
* 630301 632000: gap of unknown length
* 632302 634000: contig of 10668 bp in length
* 634303 636000: gap of unknown length
* 636304 638000: contig of 13932 bp in length
* 638305 640000: gap of unknown length
* 640306 642000: contig of 15937 bp in length
* 642307 644000: gap of unknown length
* 644308 646000: contig of 13988 bp in length
* 646309 648000: gap of unknown length
* 648310 650000: contig of 10668 bp in length
* 650311 652000: gap of unknown length
* 652312 654000: contig of 13932 bp in length
* 654313 656000: gap of unknown length
* 656314 658000: contig of 15937 bp in length
* 658315 660000: gap of unknown length
* 660316 662000: contig of 13988 bp in length
* 662317 664000: gap of unknown length
* 664318 666000: contig of 10668 bp in length
* 666319 668000: gap of unknown length
* 668320 670000: contig of 13932 bp in length
* 670321 672000: gap of unknown length
* 672322 674000: contig of 15937 bp in length
* 674323 676000: gap of unknown length
* 676324 678000: contig of 13988 bp in length
* 678325 680000: gap of unknown length
* 680326 682000: contig of 10668 bp in length
* 682327 684000: gap of unknown length
* 684328 686000: contig of 13932 bp in length
* 686329 688000: gap of unknown length
* 688330 690000: contig of 15937 bp in length
* 690331 692000: gap of unknown length
* 692332 694000: contig of 13988 bp in length
* 694333 696000: gap of unknown length
* 696334 698000: contig of 10668 bp in length
* 698335 700000: gap of unknown length
* 700336 702000: contig of 13932 bp in length
* 702337 704000: gap of unknown length
* 704338 706000: contig of 15937 bp in length
* 706339 708000: gap of unknown length
* 708340 710000: contig of 13988 bp in length
* 710341 712000: gap of unknown length
* 712342 714000: contig of 10668 bp in length
* 714343 716000: gap of unknown length
* 716344 718000: contig of 13932 bp in length
* 718345 720000: gap of unknown length
* 720346 722000: contig of 15937 bp in length
* 722347 724000: gap of unknown length
* 724348 726000: contig of 13988 bp in length
* 726349 728000: gap of unknown length
* 728350 730000: contig of 10668 bp in length
* 730351 732000: gap of unknown length
* 732352 734000: contig of 13932 bp in length
* 734353 736000: gap of unknown length
* 736354 738000: contig of 15937 bp in length
* 738355 740000: gap of unknown length
* 740356 742000: contig of 13988 bp in length
* 742357 744000: gap of unknown length
* 744358 746000: contig of 10668 bp in length
* 746359 748000: gap of unknown length
* 748360 750000: contig of 13932 bp in length
* 750361 752000: gap of unknown length
* 752362 754000: contig of 15937 bp in length
* 754363 756000: gap of unknown length
* 756364 758000: contig of 13988 bp in length
* 758365 7

BASE COUNT 54102 a 31180 c 31320 g 54266 t 80 others
ORIGIN

Query Match 1.3%; Score 40; DB 72; Length 170948;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 25113 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAA 25152

RESULT 31

AC009731 172749 bp DNA PRI 02-DEC-1999
LOCUS Homo sapiens 12 BAC RP11-438N16 (Roswell Park Cancer Institute
DEFINITION Human BAC Library) complete sequence.

AC009731

VERSION AC009731.6 GI:6492472

KEYWORDS HTG.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 172749)

REFERENCE
AUTHORS Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
Holloway,C., Hosak,H., Issar,A., Jackson,L.E., Jackson,L., Jia,Y.,
Jones,M., Kelly,S., Kneitz,S., Kondejewski,N., Kong,Y., Kovar,C.,
Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,W.,
Logan,O., Lozado,R.J., Lu,J., Lucier,R., Marondel,I., Martin,R.,
Martinez,C., McLeod,M.P., Mei,G., Merscher,S., Miller,A.,
Montgomery,K.T., Morgan,M., Morris,S., Nash,S., Nelson,A.,
Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S.,
Payton,B., Perez,L., Pu,L.L., Quiles,M., Retter,D., Rives,M.,
Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Shim,C., Simon,M.,
Sparks,A., Stamps,A., Sucgang,R., Tabor,P., Taylor,T., Vasquez,L.,
Vinson,R., Vo,Q., Wahab,M., Watlington,S., Weinstein,G.,
Weinstock,I.R., Williamson,A., Worley,K., Wren,J., Wrensford,G.,
Xiang,A.M., Yang,R., Yu,W., Zhou,X., Kucherlapati,R., Nelson,D. and
Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 172749)

AUTHORS Worley,K.C.

JOURNAL Direct Submission

TITLE Submitted (30-AUG-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 172749).

AUTHORS Worley,K.C.

JOURNAL Direct Submission

TITLE Submitted (01-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 172749)

AUTHORS Worley,K.C.

JOURNAL Direct Submission

TITLE Submitted (02-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT On Dec 1, 1999 this sequence version replaced gi:5870175.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the

entire insert of this clone. Overlapping regions of clones are only

sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 172749
Phrap values in estimate: 170704
Average error rate (BCM-Phrap estimate): 0.000441707
Fraction of Phrap values less than 40 : 30
Number of consensus changing edits: 30
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
16722 tttttgaga(n)ggagcttgc tttttgaga(t)ggagcttgc
16762 tgcagtgtg(n)gatctcgctt tgcagtgtg(c)gatctcgctt
16787 gcaacctcac(n)cccggttta gcaacctcac(t)cccggttta
19892 actagagag(n)cgatgaccc actagagag(c)cgatgaccc
23304 gcaaataggag(n)ggatgatggt gcaaataggag(g)ggatgatggt
24841 gccctgggtgg(n)aggggctagt gccctgggtgg(g)aggggctagt
24856 gctagtgtt(n)aaagtgcagt gctagtgtt(g)aaagtgcagt
24958 aaagagtatg(n)nnaaatgcta aaagagtatg(g)aaaatgcta
24959 agagtatgn(n)naaatgctaa agagtatgg(a)aaatgctaa
24960 agagtatgn(n)aaatgctaaa agagtatga(a)aaatgctaaa
31414 ggctggggggg(n)tgacccccc ggctggggggg(c)tgacccccc
31583 ggctggccgg(n)cggggggggt ggctggccgg(g)cggggggggt
31588 gccgncggg(n)ggntgaccc gccgncggg(g)ggntgaccc
31592 gncgggggg(n)tgacccccc gncgggggg(c)tgacccccc
65790 attgataaga(n)aaaaatagct attgataaga(a)aaaaatagct
67402 ataagcattg(n)ttgggacagc ataagcattg(c)ttgggacagc
73762 ggcatagccc(n)cccccccc ttccatagccc(t)cccccccc
86111 ttccataccc(n)ttcttttt ttccataccc(t)ttcttttt
86115 taacctctc(n)ttcttttt taacctctc(t)ttcttttt
90092 aatctctctg(n)ctcgnctcc aatctctctg(c)ctcgnctcc
90097 tctgntctg(n)ctcgcacaa tctgntctg(t)ctcgcacaa
90285 cccaccaga(n)tcctctccc cccaccaga(c)tcctctccc
111002 aatgctggca(n)caaaatttcc aatgctggca(a)caaaatttcc
112201 aaaaaaaa(c)caaaancaga aaaaaaaa(a)caaaancaga
112207 caaaacaaa(n)cagaatctg caaaacaaa(a)cagaatctg
112710 caaagacaa(n)caagctattt caaagacaa(a)caagctattt
160933 aagattcaga(n)ntgcccagata aagattcaga(c)ntgcccagata

* 6004 7129: contig of 1126 bp in length
* gap of unknown length
* 7130 8428: contig of 1299 bp in length
* gap of unknown length
* 8429 10151: contig of 1723 bp in length
* gap of unknown length
* 10152 10364: contig of 213 bp in length
* gap of unknown length
* 10365 11828: contig of 1464 bp in length
* gap of unknown length
* 11829 13207: contig of 1379 bp in length
* gap of unknown length
* 13208 14794: contig of 1587 bp in length
* gap of unknown length
* 14795 16552: contig of 1758 bp in length
* gap of unknown length
* 16553 18749: contig of 2197 bp in length
* gap of unknown length
* 18750 21181: contig of 2432 bp in length
* gap of unknown length
* 21182 23285: contig of 2104 bp in length
* gap of unknown length
* 23286 25187: contig of 1902 bp in length
* gap of unknown length
* 25188 27119: contig of 1932 bp in length
* gap of unknown length
* 27120 29572: contig of 2453 bp in length
* gap of unknown length
* 29573 31887: contig of 2315 bp in length
* gap of unknown length
* 31888 34036: contig of 2149 bp in length
* gap of unknown length
* 34037 35754: contig of 1718 bp in length
* gap of unknown length
* 35755 38353: contig of 2599 bp in length
* gap of unknown length
* 38354 41887: contig of 3534 bp in length
* gap of unknown length
* 41888 44323: contig of 2436 bp in length
* gap of unknown length
* 44324 46990: contig of 2667 bp in length
* gap of unknown length
* 46991 50211: contig of 3221 bp in length
* gap of unknown length
* 50212 53782: contig of 3571 bp in length
* gap of unknown length
* 53783 57135: contig of 3353 bp in length
* gap of unknown length
* 57136 60737: contig of 3602 bp in length
* gap of unknown length
* 60738 64425: contig of 3688 bp in length
* gap of unknown length
* 64426 69302: contig of 4877 bp in length
* gap of unknown length
* 69303 74567: contig of 5265 bp in length
* gap of unknown length
* 74568 79311: contig of 4744 bp in length
* gap of unknown length
* 79312 82695: contig of 3384 bp in length
* gap of unknown length
* 82696 88579: contig of 5884 bp in length
* gap of unknown length
* 88580 93341: contig of 4762 bp in length
* gap of unknown length
* 93342 97484: contig of 4143 bp in length
* gap of unknown length
* 97485 101749: contig of 4265 bp in length
* gap of unknown length
* 101750 108833: contig of 7084 bp in length
* gap of unknown length
* 108834 114784: contig of 5951 bp in length
* gap of unknown length
* 114785 123180: contig of 8396 bp in length

* 123181 129628: contig of 6448 bp in length
* gap of unknown length
* 129629 137374: contig of 7746 bp in length
* gap of unknown length
* 137375 145275: contig of 7901 bp in length
* gap of unknown length
* 145276 153859: contig of 8584 bp in length
* gap of unknown length
* 153860 165597: contig of 11738 bp in length
* gap of unknown length
* 165598 176771: contig of 11174 bp in length
* gap of unknown length
* 176772 193632: contig of 16861 bp in length.

FEATURES

source

1. .193632
/organism="Homo sapiens"
/db_xref="taxon:9606"
/Chromosome="15"
/map="15"
/clone="RP11-344C1"
/clone_lib="RP11-344C1" Human Male BAC"
BASE COUNT 57762 a 38404 c 38424 g 59032 t 10 others
ORIGIN

Query Match 1.3%; Score 40; DB 43; Length 193632;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 55291 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 55252

RESULT 33

HS80N2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT


```

/note="AluJo repeat: matches 122. .299 of consensus"
42029. .42097
/note="L1MC/D repeat: matches 5752. .5819 of consensus"
42167. 42204
/note="19 copies 2 mer gt 89% conserved"
42170. .42205
/note="9 copies 4 mer tgtg 88% conserved"
43073. 43153
/note="L2 repeat: matches 2067. .2146 of consensus"
43193. 43528
/note="THE1B repeat: matches 1. .364 of consensus"
43600. 43907
/note="AluJo repeat: matches 5. .310 of consensus"
44113. 44399
/note="match: STS: Em:Z79002"
44859. 44900
/note="L1M3c repeat: matches 273. .314 of consensus"
44915. 45163

Query Match 1.3%; Score 40; DB 11; Length 193731;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tcacatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 63526 TCACATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 63565

RESULT 34
BTCIB8
LOCUS B.taurus CI-B8 mRNA 583 bp MAM 27-AUG-1992
DEFINITION B.taurus CI-B8 mRNA for ubiquinone oxidoreductase complex.
ACCESSION X63219
VERSION X63219.1 GI:245
KEYWORDS NADH dehydrogenase; NADH-ubiquinone oxidoreductase complex.
SOURCE Bos taurus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 583)
AUTHORS Walker,J.
TITLE Submitted (12-NOV-1991) Walker J., MRC Lab. of Molecular Biology,
JOURNAL Hills Road, Cambridge CB2 2QH, England
REFERENCE 2 (bases 1 to 583)
AUTHORS Walker,J.E., Arizumendi,J.M., Dupuis,A., Fearnley,I.M., Finel,M.,
Medd,S.M., Pilkington,S.J., Runswick,M.J. and Skehel,J.M.
TITLE Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from
bovine heart mitochondria. Application of a novel strategy for
sequencing proteins using the polymerase chain reaction
JOURNAL J. Mol. Biol. 226 (4), 1051-1072 (1992)
MEDLINE 92389317
FEATURES
Location/Qualifiers
source
1. .583
/organism="Bos taurus"
/db_xref="taxon:9913"
/germline
/tissue_type="heart"
34. .333
/gene="CI-B8"
34. .333
/gene="CI-B8"
/EC_number="1.6.99.3"
/codon_start=1
/evidence=experimental
/product="NADH dehydrogenase"
/protein_id="CAA44904.1"
/db_xref="GI:246"
/db_xref="SWISS-PROT:Q02370"
/translation="MAAAATRGVKGKGLRIRIHLQSRPGSGVDRFIEKRYVEL
KKANPLDILIRECSDVQPKLWARYAFQGEKNVSLNFSADQVTRALENVLSKA"
polyA_signal 523. .528
polyA_site 551

```

```

BASE COUNT 183 a 126 c 143 g 131 t
ORIGIN
Query Match 1.3%; Score 39; DB 3; Length 583;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 545 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 583

RESULT 35
AF206698
LOCUS Epiblena scudderiana muscle LIM protein (Mlp) mRNA 606 bp INV 09-JAN-2000
DEFINITION Epiblena scudderiana muscle LIM protein (Mlp) mRNA, complete cds.
ACCESSION AF206698
VERSION AF206698.1 GI:6683944
KEYWORDS Epiblena scudderiana.
SOURCE Epiblena scudderiana.
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Eukaryota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Tortricodea; Tortricidae; Olethreutinae; Epiblena.
REFERENCE 1 (bases 1 to 606)
AUTHORS Bilgen,T., English,T.E. and Storey,K.B.
TITLE EsMlp, a muscle-LIM protein gene, is up-regulated during cold
exposure in the freeze-avoiding larvae of Epiblena scudderiana
JOURNAL Unpublished
AUTHORS Bilgen,T. and Storey,K.B.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1999) Biology, Carleton University, 1125 Colonel
By Drive, Ottawa, ON K1S 5B6, Canada
FEATURES
Location/Qualifiers
source
1. .606
/organism="Epiblena scudderiana"
/db_xref="taxon:111441"
/note="Induced by subzero exposure; highest expression
apparently localized in pharyngeal muscles"
1. .606
/gene="Mlp"
93. 377
/gene="Mlp"
/codon_start=1
/product="muscle LIM protein"
/protein_id="AAF23406.1"
/db_xref="GI:6683945"
/translation="MPFKPADNPKPCGKSVAAERVAGGLKWHKTCFKGCMCNKS
LDSTNCTEHDGEIYCKNCHGRFPGKYGFGGAGCLSDAGAQFQENAN"
BASE COUNT 192 a 153 c 141 g 120 t
ORIGIN
Query Match 1.3%; Score 39; DB 34; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 565 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 603

RESULT 36
AF123275
LOCUS Dictyostellium discoideum superoxide-generating NADPH oxidase
DEFINITION Dictyostellium discoideum superoxide-generating NADPH oxidase
ACCESSION AF123275
VERSION AF123275.1 GI:4530485
KEYWORDS

```

```

CDs      83. .2092
/partial
/codon_start=1
/product="Na+ channel"
/protein_id="AAA21813.1"
/db_xref="GI:493605"
/translation="MEGNKLEEQSDSPQTLGMLKGNKREOGLGPEPAPAAQPOPTAE
EEALIEFHRSYRELFEFCNNNTTHGAIRLVCSOHRMKTAFWALWICLTCGMYYWQF
GLLFGEYFVSFVSINILNLSDKLVFPVATICTLNPYRYPEIKKEELEDRITEOTLFD
LYKVSFTTILVAGSRSDRLGTLPHLQRLRVPPPHGARRARSVASSLRDNNPQVD
WKKDKIGFOLCQNKQDCPQVQYSSGVDAVREWHYINILSRLETLSLEEDTLG
NFTACRFNVQSCQANYSHFHPHYNCYTFDKNNSNLMSSMPGGINGLSLMLRA
EQNDFIPLLSTVTGAVRHVGQDEPAMDGGFNLPGVETSISMKEITLDRLGDDY
DCTKNSDVENVNLYPSKYTQQVCIHSCFQESMIKEGCAYIFYPRQNVYCYDKRKH
SSWGYCYIKLQVDFSSDLHGCTKCRKPCSVTSYQLSAGYSRMPSVTSQEWVFWMLSR
SNWYTVNNKRGVAKVNIYFELKELNYKTNSPVSVMYTLNLSNGSOWSLWFGSSVLSV
VENAEVLFDLLVIMFLMLRRFRSRYSPGSGRGAEVASTLASSPSHFCPIHMSL
SLSQGGPAPSPALTPAPPATAYATGPRFSPGSGAGASSSTCPLGP"
BASE COUNT      486 a 739 c 632 g 477 t
ORIGIN

Query Match      1.38; Score 39; DB 9; Length 2334;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcgtccaacctggcagcagctgagcctgtggtt 1838
|||||
DB 1720 CCTCCTGTCACCTGGCAGCCAGTGAGCTGTGGTT 1758
|||||

RESULT 38
HSAENAC4      2562 bp DNA PRI 02-MAY-1999
LOCUS Homo sapiens epithelial sodium channel alpha-subunit gene, exons 9
DEFINITION through 13 and complete cds.
ACCESSION AF060913
VERSION AF060913.1 GI:4731107
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2562)
AUTHORS Chow Y.H., Wang Y., Plumb J., O'Brodovich H. and Hu J.
TITLE Hormonal regulation and genomic organization of the human
amiloride-sensitive epithelial sodium channel alpha-subunit gene
JOURNAL Pediatr. Res. (1999) In press
REFERENCE 2 (bases 1 to 2562)
AUTHORS Chow Y.H. and Hu J.
TITLE Direct Submission
JOURNAL Submitted (22-APR-1998) Lung Biology, Hospital for Sick Children,
555 University Ave, Toronto, ON M5G 1X8, Canada
FEATURES
source
1. .2562
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13"
join(AF060910.1:3085..3129,AF060910.1:3797..4266,
AF060911.1:1675..1942,AF060911.1:3781..3971,
AF060912.1:728..831,AF060912.1:1173..1336,
AF060912.1:1760..1858,AF060912.1:2053..2170,20..98,
204..261,420..475,623..698,1172..2562)
/product="epithelial sodium channel alpha-subunit"
join(AF060910.1:3851..4266,AF060911.1:1675..1942,
AF060911.1:3781..3971,AF060912.1:728..831,
AF060912.1:1173..1336,AF060912.1:1760..1858,
AF060912.1:2053..2170,20..98,204..261,420..475,623..698,
1172..1552)
/note="alpha-hENaC"
/codon_start=1

```

```
/product="epithelial sodium channel alpha-subunit"
/protein_id="AAD28355.1"
/db_xref="GI:47311109"
/translation="MEGNKLEQDSSPQSTPGLMKGNKREQGLGPPEAPAAQOQTAE
EALIEFHRSYRELFEFFCNNTIIGAIRLVCQSHNRMTAFWVWLCTFGMMYQWF
GLUFGESYSPVSLNINLSNGLKLVFPVAVTICTLNRYPEIKEELEELDRITEQTLFD
LYKYSFTTLVAGSRSRDLRGLTLPQLRVRPPPHGARRARSVASLRDNNQVOD
WKMDKIFGOLCQNKSDCFYQYSSGVDAVREWFHYINILSRLEPTLPSLEEDTLG
NFIACRNOVSCNOANYSHFHHPMGVNCYTFENDKNSNLWSSMPGINNGLSLMLRA
EONDRITPLLSTVIGARVMVHGOEPAFMDGDFGNLRPGVETISMRKETLDRLGDDYG
DCTKNGSDVPVENLPSKYTQQVCIHSCFQESMIKECCGAYIFYPRQNVQYCDYRKH
SWGKCYIQLQVDFSDHLGCTCKRKCPSVTSYQLSAGISRWPSVTSQEWYFQMLSR
QNNYVNNKRGVAKVNIFFKELNKTNSPESVMTVTLNLSGQSWLWFGSSVLSV
VMAELVFDLLVIMFLMLLRFRSRYWSPGRGCGAQEAVSTLASSPPSHFCHPMSL
SLSQGPAPSPALTAPPAYATLGRPSPGSGAGSSSTCPLGGP"
exon
20. .98
/number=9
exon
204. .261
/number=10
exon
420. .475
/number=11
exon
623. .698
/number=12
exon
1172. .2562
/number=13
BASE COUNT 533 a 747 c 689 g 593 t
ORIGIN
Query Match 1.3%; Score 39; DB 39; Length 2562;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1800 cctcctgtccaaacctggcagcagtgagcctgtggtt 1838
|||||
Db 1180 CCTCCTGTCCAACTGGCGCAGTGAGCCTGTGTT 1218
|||||

RESULT 39
HSLASNA HSLASNA 3151 bp mRNA PRI 01-FEB-1994
LOCUS H.sapiens mRNA for lung amiloride sensitive Na+ channel protein.
DEFINITION X76180
ACCESSION X76180
VERSION 1
KEYWORDS Na+ channel; Na+ channel protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3151)
AUTHORS Barbry,P.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1993) P. Barbry, CNRS, IPCM 650 Route des
Lucioles, 06560 Sophia Antipolis, FRANCE
REFERENCE 2 (bases 1 to 3151)
AUTHORS Voilley,N., Lingueglia,E., Champigny,G., Mattei,M.G., Waldmann,R.,
Lazdunski,M. and Barbry,P.
TITLE The lung amiloride-sensitive Na+ channel: biophysical properties,
pharmacology, ontogenesis, and molecular cloning
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (1), 247-251 (1994)
MEDLINE 94105144
FEATURES Location/Qualifiers
source
1. .3151
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/tissue_type="lung"
/cell_type="epithelial"
/dev_stage="adult"
100. .2109
/codon_start=1
/product="Na+ channel protein"
/protein_id="CAA53773.1"
CDS
```

```
/db_xref="GI:452650"
/translation="MEGNKLEQDSSPQSTPGLMKGNKREQGLGPPEAPAAQOQTAE
EALIEFHRSYRELFEFFCNNTIIGAIRLVCQSHNRMTAFWVWLCTFGMMYQWF
GLUFGESYSPVSLNINLSNGLKLVFPVAVTICTLNRYPEIKEELEELDRITEQTLFD
LYKYSFTTLVAGSRSRDLRGLTLPQLRVRPPPHGARRARSVASLRDNNQVOD
WKMDKIFGOLCQNKSDCFYQYSSGVDAVREWFHYINILSRLEPTLPSLEEDTLG
NFIACRNOVSCNOANYSHFHHPMGVNCYTFENDKNSNLWSSMPGINNGLSLMLRA
EONDRITPLLSTVIGARVMVHGOEPAFMDGDFGNLRPGVETISMRKETLDRLGDDYG
DCTKNGSDVPVENLPSKYTQQVCIHSCFQESMIKECCGAYIFYPRQNVQYCDYRKH
SWGKCYIQLQVDFSDHLGCTCKRKCPSVTSYQLSAGISRWPSVTSQEWYFQMLSR
QNNYVNNKRGVAKVNIFFKELNKTNSPESVMTVTLNLSGQSWLWFGSSVLSV
VMAELVFDLLVIMFLMLLRFRSRYWSPGRGCGAQEAVSTLASSPPSHFCHPMSL
SLSQGPAPSPALTAPPAYATLGRPSPGSGAGSSSTCPLGGP"
BASE COUNT 677 a 995 c 803 g 676 t
ORIGIN
Query Match 1.3%; Score 39; DB 9; Length 3151;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1800 cctcctgtccaaacctggcagcagtgagcctgtggtt 1838
|||||
Db 1737 CCTCCTGTCCAACTGGCGCAGTGAGCCTGTGTT 1775
|||||

RESULT 40
A39975 A39975 4028 bp DNA PAT 05-MAR-1997
LOCUS Sequence 8 from Patent WO9421790.
DEFINITION A39975
ACCESSION A39975
VERSION A39975.1 GI:2296171
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4028)
AUTHORS Hayden,M., Goldberg,P., Andrew,S., Rommens,J.M., Lin and Blaoyang,
TITLE PROCESS FOR ISOLATING GENES AND THE GENE CAUSATIVE OF HUNTINGTON'S
DISEASE AND DIFFERENTIAL 3' POLYADENYLATION IN THE GENE
JOURNAL Patent: WO 9421790-A 8 29-SEP-1994;
COMMENT UNIV BRITISH COLUMBIA (CA)
Other publication CA 2092455 940926
Other publication AU 6281094 941011
Other publication US 5534438 960709
Other publication CA 2158918 940929.
FEATURES Location/Qualifiers
source
1. .4028
/organism="unidentified"
/db_xref="taxon:32644"
/chromosome="4"
BASE COUNT 838 a 1100 c 1140 g 950 t
ORIGIN
Query Match 1.3%; Score 39; DB 5; Length 4028;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 3984 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4022
|||||

RESULT 41
LOCUS I23502
DEFINITION I23502
ACCESSION I23502
VERSION I23502.1 GI:1603372
KEYWORDS Unknown.
SOURCE I23502
I23502 4032 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 8 from patent US 5534438.
ACCESSION I23502
VERSION I23502.1 GI:1603372
KEYWORDS Unknown.
SOURCE
```


/usedin-z92978:SCNNIA_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"

BASE COUNT 1326 a 1778 c 1639 g 1570 t
ORIGIN

Query Match 1.3%; Score 39; DB 10; Length 6313;
Best Local Similarity 100.0%; Pred. No. 3.1e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccaaacctggcagccagtcggagcctgtgtgtt 1838
|||||
Db 4916 CCTCCTGTCCAACTGGCAGCCAGTCGGAGCCTGTGTGT 4954

RESULT 43

HS514K20

LOCUS

DEFINITION

HS514K20 100368 bp DNA PRI 23-NOV-1999

Human DNA sequence from PAC 514K20 on chromosome 6p22.3-24.3. EST,

CA repeats and STS.

ACCESSION

AL008731

VERSION

AL008731.1

KEYWORDS

6p22.3-24.3; repeat polymorphism.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

I. (bases 1 to 100368)

Wild.A.

Direct Submission

Submitted (09-JAN-1998) Chromosome 6 Project Group

(http://www.sanger.ac.uk/HGP/Chr6/) Sanger Centre, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On Jan 13, 1998 this sequence version replaced gi:2598467.

IMPORTANT: This sequence is the entire insert of clone 514K20.

During sequence assembly data are compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variations annotated may not be found in the sequence submission

corresponding to the overlapping clone as we submit sequences with

only a small overlap as described above.

This sequence was generated from part of bacterial clone contigs of

human chromosome 6, constructed by the Sanger Centre chromosome 6

mapping group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6/

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

The true left end of clone 514K20 is at 1 in this sequence. The

true right end of clone 514K20 is at 100368.

514K20 is from the library RPCI3 constructed at the Roswell Park

Cancer Institute by the group of Pieter de Jong.

For further details see <http://bacpac.med.buffalo.edu/>.

Location/Qualifiers

1..100368

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/map="6p22.3-24.3"

/clone="RP3-514K20"

/clone_lib="RPCI-3"

239..534

repeat_region

/note="AluJo repeat: matches 1..294 of consensus"

repeat_region

1882..2174

/note="AluJo repeat: matches 2..301 of consensus"

repeat_region

3503..3808

/note="AluSx repeat: matches 2..301 of consensus"

repeat_region

4742..5056

/note="AluJo repeat: matches 1..301 of consensus"

repeat_region

5276..5897

/note="L1MB3 repeat: matches 291..923 of consensus"

repeat_region

6167..6299

/note="L1MB3 repeat: matches 767..908 of consensus"

repeat_region

9707..9975

/note="MLT1B repeat: matches 390..113 of consensus"

repeat_region

10357..10552

/note="MLT1C repeat: matches 466..268 of consensus"

repeat_region

10571..11046

/note="MER4A2 repeat: matches 1..505 of consensus"

repeat_region

11042..11286

/note="MLT1C repeat: matches 247..1 of consensus"

repeat_region

12252..12370

/note="MLT1B repeat: matches 5..120 of consensus"

repeat_region

14219..14506

/note="match: 116457 STS containing (CA) repeat"

repeat_region

14279..14319

/note="match: 20 copies of CA 100 % conserved; differs from

216457; D6S259"

repeat_region

15184..15235

/note="2 copies of 26 mer 98 % conserved"

repeat_region

18638..18929

/note="AluSx repeat: matches 302..6 of consensus"

repeat_region

19749..20135

/note="MLT1C repeat: matches 23..466 of consensus"

repeat_region

21690..21761

/note="MIR repeat: matches 186..256 of consensus"

repeat_region

23026..23327

/note="AluSx repeat: matches 1..302 of consensus"

repeat_region

23328..23478

/note="AluJb repeat: matches 122..273 of consensus;

incomplete repeat"

repeat_region

24403..24786

/note="L1MB2 repeat: matches 469..884 of consensus"

repeat_region

24779..24937

/note="MER4A2B repeat: matches 1149..1300 of consensus"

repeat_region

26000..26282

/note="AluJb repeat: matches 293..1 of consensus"

repeat_region

29619..29921

/note="AluJb repeat: matches 2..301 of consensus"

repeat_region

31380..31670

/note="AluSx repeat: matches 303..1 of consensus"

repeat_region

31687..31753

/note="MIR2 repeat: matches 141..75 of consensus"

repeat_region

31926..32455
/note="L1MB8 repeat: matches 1035..514 of consensus"
32475..32737
/note="L1MB1 repeat: matches 566..302 of consensus"
32859..33179
/note="AluJo repeat: matches 302..1 of consensus"
33831..33913
/note="MLT1C repeat: matches 89..1 of consensus"
34014..34307
/note="AluJb repeat: matches 1..298 of consensus"
35042..35117
/note="MIR2 repeat: matches 146..74 of consensus"
35186..35487
/note="AluSx repeat: matches 1..303 of consensus"
36019..36447
/note="MLT1C repeat: matches 466..1 of consensus"
38175..38477
/note="AluSx repeat: matches 302..1 of consensus"
38520..38643
/note="MIR2 repeat: matches 145..16 of consensus"
39015..39237
/note="L1MB1 repeat: matches 432..661 of consensus"
39457..39631

AUTHORS
TITLE
JOURNAL

Pavitt,R.
Direct Submission
Submitted (19-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerquests@sanger.ac.uk
On Apr 21, 2000 this sequence version replaced gi:7160601.
----- Genome Center

COMMENT

Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba93B10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid, L08752; 100% of reads
Chemistry: Dye-terminator ET-amersham; 14% of reads Chemistry:
Dye-terminator Big Dye; 85% of reads
Consensus quality: 140176 bases at least Q40
Consensus quality: 146899 bases at least Q30
Consensus quality: 151390 bases at least Q20
Insert size: 154627; sum-of-contigs
Insert size: 125380; 18.5% error; agarose-fp
Quality coverage: 2.97x in Q20 bases; sum-of-contigs Quality
coverage: 3.67x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently * consists
of 27 contigs. The true order of the pieces is * not known and
their order in this sequence record is * arbitrary. Where the
contigs adjacent to the vector can * be identified, they are the
labelled with 'clone.end' in the * feature table. Some order and
orientation information * can tentatively be deduced from paired
sequencing reads * which have been identified to span the gap
between two * contigs. These are labelled as part of the same *
'fragment_chain', and the order and relative orientation * of the
pieces within a fragment_chain is reflected in * this file. Gaps
between the contigs are represented as * runs of N, but the exact
sizes of the gaps are unknown. * This record will be updated with
the finished sequence as * soon as it is available and the
accession number will be * preserved.
*
1 9746 contig of 9746 bp in length; fragment_chain 1 *
9847 29417 contig of 19571 bp in length; fragment_chain 1 *
29518 37118 contig of 7601 bp in length; fragment_chain 1 *
37219 38248 contig of 1030 bp in length; fragment_chain 2 *
38349 48422 contig of 10074 bp in length; fragment_chain 2 *
48523 62332 contig of 13811 bp in length; fragment_chain 2 *
62434 59691 contig of 7258 bp in length; fragment_chain 2 *
69792 71246 contig of 1455 bp in length; fragment_chain 2 *
71347 73427 contig of 2081 bp in length; fragment_chain 2 *
73528 76614 contig of 3087 bp in length; fragment_chain 3 *
76715 80354 contig of 3640 bp in length; fragment_chain 3 *
80455 82645 contig of 2191 bp in length; fragment_chain 3 *
82746 91544 contig of 8799 bp in length; fragment_chain 3 *
91645 93404 contig of 1760 bp in length; fragment_chain 3 *
93505 95394 contig of 1890 bp in length; fragment_chain 4 *
95495 98525 contig of 3031 bp in length; fragment_chain 4 *
98626 104361 contig of 5736 bp in length; fragment_chain 4 *
104462 109443 contig of 4982 bp in length; fragment_chain 4 *
109544 125125 contig of 15582 bp in length; fragment_chain 5 *
125226 127995 contig of 2770 bp in length; fragment_chain 5 *
128096 130614 contig of 2519 bp in length;
* 130715 133455 contig of 2741 bp in length
* 133556 137090 contig of 3535 bp in length
* 137191 138724 contig of 1534 bp in length
* 138825 150575 contig of 11751 bp in length; fragment_chain 6 *
* 150676 152322 contig of 1647 bp in length; fragment_chain 6 *
152423 157227 contig of 4805 bp in length; fragment_chain 6.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
Location/Qualifiers
1. .157227

FEATURES
source

misc_feature
1. .9746
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q12-13.1"
/clone="RP11-93B10"
/clone_lib="RPC1-11.1"
/note="assembly_fragment:00809
fragment_chain:1
clone_end:r7
vector_side:left"
9847. .29417
/note="assembly_fragment:00615
fragment_chain:1"
29518. 37118
/note="assembly_fragment:00573
fragment_chain:1"
37219. 38248
/note="assembly_fragment:00774
fragment_chain:2"
38349. 48422
/note="assembly_fragment:00644
fragment_chain:2"
48523. 62333
/note="assembly_fragment:00969
fragment_chain:2"
62434. 69691
/note="assembly_fragment:00900
fragment_chain:2"
69792. 71246
/note="assembly_fragment:00672
fragment_chain:2"
71347. 73427
/note="assembly_fragment:00184
fragment_chain:2"
73528. 76614
/note="assembly_fragment:00955
fragment_chain:3"
76715. 80354
/note="assembly_fragment:00440
fragment_chain:3"
80455. 82645
/note="assembly_fragment:00594
fragment_chain:3"
82746. 91544
/note="assembly_fragment:00816
fragment_chain:3"
91645. 93404
/note="assembly_fragment:00707
fragment_chain:3"
93505. 95394
/note="assembly_fragment:00151
fragment_chain:4"
95495. 98525
/note="assembly_fragment:00098
fragment_chain:4"
98626. 104361
/note="assembly_fragment:00365
fragment_chain:4"
104462. 109443
/note="assembly_fragment:00981
fragment_chain:4"
109544. 125125
/note="assembly_fragment:00602
fragment_chain:5"
125226. 127995
/note="assembly_fragment:00818
fragment_chain:5"
128096. 130614
/note="assembly_fragment:00295"
130715. 133455
/note="assembly_fragment:00665"
133556. 137090


```
misc_feature /note="assembly_fragment:00849"
137191. .138724
misc_feature /note="assembly_fragment:01034"
138825. .150575
/note="assembly_fragment:00274
fragment_chain:6"
misc_feature /note="assembly_fragment:00558
150676. .152322
fragment_chain:6"
misc_feature /note="assembly_fragment:00317
152423. .157227
fragment_chain:6
clone_end:SP6
vector_side:right"
BASE COUNT 49599 a 30099 c 29258 g 45662 t 2609 others
ORIGIN
```

```
Query Match 1.3%; Score 39; DB 40; Length 157227;
Best Local Similarity 100.0%; Pred. No. 4.3e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3115
|||||
Db 97866 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 97904
|||||
```

Search completed: September 11, 2000, 01:02:03
Job time: 1111 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:26:52 ; Search time 2749.07 Seconds
(without alignments)
2023.556 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcgaccacgcgtcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues
Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: em_fun:*
17: em_hum1:*
18: em_hum2:*
19: em_in:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_sy:*
29: em_un:*
30: em_v1:*
31: gb_htg1:*
32: gb_htg2:*
33: gb_in1:*
34: gb_in2:*
35: em_ba1:*
36: em_ba2:*
37: em_hum3:*
38: em_hum4:*
39: gb_pr4:*
40: gb_htg3:*
41: gb_htg4:*
42: gb_htg5:*
43: gb_htg6:*

- 44: gb_htg7:*
45: em_htg1:*
46: em_htg2:*
47: em_htg3:*
48: em_hum5:*
49: gb_pl3:*
50: gb_pr5:*
51: gb_htg8:*
52: gb_htg9:*
53: gb_htg10:*
54: gb_htg11:*
55: gb_htg12:*
56: gb_htg13:*
57: gb_htg14:*
58: gb_in3:*
59: gb_htg15:*
60: gb_htg16:*
61: gb_htg17:*
62: em_htg4:*
63: em_htg5:*
64: em_htg6:*
65: em_htg7:*
66: em_hum6:*
67: gb_htg18:*
68: gb_htg19:*
69: gb_htg20:*
70: gb_htg21:*
71: gb_htg22:*
72: gb_htg23:*
73: gb_htg24:*
74: gb_htg25:*
75: gb_htg26:*
76: gb_htg27:*
77: gb_htg28:*
78: gb_htg29:*
79: gb_htg30:*
80: gb_htg31:*
81: gb_vil:*
82: gb_vil2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	3117	100.0	3117	12	RNASNA	X70497	R. norvegic
2	3062.8	98.3	3081	12	RNASNA	X70521	R. norvegic
3	2280	73.1	3000	12	AF112185	AF112185	Mus muscu
4	2196.8	70.5	2200	12	RNU54700	U54700	Rattus norv
5	2195.2	70.4	2200	12	RNU54699	U54699	Rattus norv
6	1532.2	49.2	3151	9	HSLASNA	X76180	H.sapiens m
7	1436.8	46.1	2334	9	HUMSODIUM	L29007	Human Kidne
8	1401.8	45.0	3586	3	BTU14944	U14944	Bos taurus
9	1375	44.1	2695	12	CP0249296	AJ249296	Cavia por
10	1362.2	43.7	2915	3	OCU132108	AJ132108	Oryctolag
11	1357	43.5	3385	12	AF071230	AF071230	Cavia cob
12	752.2	24.1	3368	4	GGU58475	U58475	Gallus gall
13	751.8	24.1	1003	3	AF232715	AF232715	Ovis arie
14	747.4	24.0	2321	4	GGU62902	U62902	Gallus gall
15	747.4	24.0	3399	4	GGU62903	U62903	Gallus gall
16	729.8	23.4	1065	3	AF229025	AF229025	Oryctolag
17	584.6	18.8	2773	4	XL023535	U23535	Xenopus lae
18	502	16.1	3841	4	GGU62904	U62904	Gallus gall
19	318.4	10.2	1202	10	HSZ92978	Z92978	H.sapiens S
20	318.4	10.2	4526	39	HSANAC1	AF060910	Homo sapi
21	318.4	10.2	140026	41	AC005840	AC005840	Homo sapi
22	313.4	10.1	188488	39	AC006057	AC006057	Homo sapi
23	307.6	9.9	1917	11	HSU38254	U38254	Human amilo
24	305.8	9.8	1917	11	AF038165	AF038165	Pan trogl

25	303.4	9.7	6313	10	HS292981	252981 H. sapiens S
26	303	9.7	2562	39	HSANAC4	AF060913 Homo sapi
27	295.6	9.5	2448	3	OCU132110	AJ132110 Oryctolag
28	293	9.4	344	12	MWU52006	U52006 Mus musculu
29	290.6	9.3	2136	12	RNU37539	U37539 Rattus norv
30	290.6	9.3	3012	12	RNENACG	X77933 R.norvegicu
31	288.4	9.3	1923	10	HUMBETA	L36593 Homo sapien
32	287.4	9.2	2136	12	RNU37540	U37540 Rattus norv
33	287.4	9.2	2938	12	RNNAC	X78034 R.norvegicu
34	286.8	9.2	2564	9	HSCNN1B	X87159 H.sapiens m
35	284	9.1	2436	12	AF112186	AF112186 Mus muscu
36	283.6	9.1	749	12	AF082073	AF082073 Rattus no
37	283.6	9.1	2223	12	AF081783	AF081783 Rattus no
38	282.8	9.1	2989	12	AF112187	AF112187 Mus muscu
39	278.4	8.9	2221	12	RNU35174	U35174 Rattus norv
40	276.8	8.9	2221	12	RNU35175	U35175 Rattus norv
41	276.8	8.9	2462	12	RNENACB	X77932 R.norvegicu
42	274.6	8.8	2204	10	HUMGAMM	L36592 Homo sapien
43	272.6	8.7	2495	3	OCU132109	AJ132109 Oryctolag
44	271.4	8.7	3384	9	HSCNN1G	X87160 H.sapiens m
45	262.2	8.4	3022	4	XLU25342	U25342 Xenopus lae
ALIGNMENTS						
RESULT 1						
RENACA	RNENACA	3117 bp	mRNA	ROD	08-MAR-1994	
LOCUS	R. norvegicus mRNA for sodium channel, alpha subunit.					
DEFINITION	X70497					
ACCESSION	X70497.1 GI:458845					
VERSION	alpha subunit; sodium channel.					
KEYWORDS	Norway rat.					
SOURCE	Rattus norvegicus					
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE	1 (bases 1 to 3117)					
AUTHORS	Canessa, C.M., Horisberger, J.D. and Rossier, B.C.					
TITLE	Epithelial sodium channel related to proteins involved in neurodegeneration					
JOURNAL	Nature 361 (6411), 467-470 (1993)					
MEDLINE	93156815					
REFERENCE	2 (bases 1 to 3117)					
AUTHORS	Canessa, C.M.					
TITLE	Direct Submission					
JOURNAL	Submitted (02-MAR-1994) C.M. Canessa, Institut de Pharmacologie et Toxicologie, Bugnon 27, 1005 Lausanne, SWITZERLAND					
FEATURES	Location/Qualifiers					
source	1..3117					
	/organism="Rattus norvegicus"					
	/strain="Sprague Dawley"					
	/db_xref="taxon:10116"					
	/dev_stage="adult"					
	/tissue_type="distal colon epithelium"					
gene	82..2178					
	/gene="rENaC"					
CDS	82..2178					
	/gene="rENaC"					
	/codon_start=1					
	/product="epithelial sodium channel, alpha subunit"					
	/protein_id="CAA49905.1"					
	/db_xref="GI:458846"					
	/db_xref="SWISS-PROT:P37089"					
	/translation="MLDHTRAPELMIDLHASNPKSMKNQKEQDPCPPQMQG LGKGRREQGLGPPAPRQPTFEEALIEPHRSYRELFPQFNNTIHGAIRLVCS KINRMTAFWALVLCFTGMVWQFALLFEYLSYPVSLINLNSDKLVFPVAVCTL NPYRYTEKEELEELDRTEOTLFDLYKYNSSYTRQAGARRSSRDLLGAFPHPLQRL RPPPPYSRTARSSSSVRDNNPOVDKDWKIGFQLCNQKSDCFYOTYSSGVDAVR EWRPHYINILRSDTSPALSEEALGNFIPTFRNQAPCNQANYSKHHPMWGNCTV FNDKNNLSMWSMPGVNNGLSLTLEQNDIFPLLSVTIGARVWHGGDEPAFMDDG GFNLRPQVETSI5MRKEALDSLGNQDCTENGSDVPVKNLYPSKYTQQVCVTHSCFQE NMIKCGGAYIFVYPKGVFECDYRKQSWGVCYVKLOGAFSLDSLGCFSKRKPFCV					



|||||
Db 901 GCAGTGGAGAGTGGTACGGCTTCCATTACATCAACATTCTGTCCAGACTGTGCGACACC 960
QY 961 tcgcccgtctagagaaagccctgggcaacttctcatcttcaacctgtcgcttcaaccag 1020
Db 961 TCGCCCGCTCTAGAGAAAGGCCCTGGGCAACTTCATCTTCACTTCCACCTGTGGCTTCAACCCAG 1020
QY 1021 gccctctgaaccagggcgaattattccaagttccacacccccaatgtacggaactgtac 1080
Db 1021 GCCCCCTGCAACACAGCGGAATATTCCAAGTTCCACCACCCCATGTACGGGAACCTGCTAC 1080
QY 1081 actttcaatgcagaacaaactccaatctcttgatgtcctccaatgcctggagtcacaat 1140
Db 1081 ACTTTCAATGCAGAAGAACAACTCCAATCTCTGGATGCTCCTCCATGCTGGAGTCAACAAT 1140
QY 1141 ggttgcctcaactgcacagcgacagagaagaatgacttcaatccccctgtgtccacagt 1200
Db 1141 GGTTTGTCCCTGACATCGCGACAGAGCAGATGATCTTATCCCCCTGCTGTCCACAGTG 1200
QY 1201 acggggccaggggtgatggtcaggtcaggatgagccctgcctttatggatgatggtg 1260
Db 1201 ACGGGGCCAGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 ttaacttgagcctggcgtgagacccctccatcagtatgagaaagaaocctggacagc 1320
Db 1261 TTCAACTTGAGGCGCTGGCGTGGAGACCTCCATCAGTATGAGAAAGGAGCCCTGGACAGC 1320
QY 1321 ctcgaggaataatcagcgacgttactgagaatggtagcgaagtccccgttcaagaacctt 1380
Db 1321 CTCGGAGGAAATTTACGGCGACTGTACTGAGAAATGGTAGCGATGTCCCGGTCAAGAACCCT 1380
QY 1381 tacccttccaagtatacacagcaggtgtgtcattcaactcctgcttccagagacaatgatc 1440
Db 1381 TACCCTTTCCAGTATACACAGCAGGTGTGCATTCATCTCTGCTTCCAGGAGAACATGATC 1440
QY 1441 aagaagtgtggtgtgcctacatcttctaccctaaagggaggttgaagttctgtgac 1500
Db 1441 AAGAAGTGTGGCTGTGCCCTACATCTTCTACCCCTTAAGCCCAAGGAGTGTGAGTCTGTGAC 1500
QY 1501 taccgaagcagagctccctgggctattgctattataaactgcagggccttctccttg 1560
Db 1501 TACCGAAACAGAGCTCCCTGGGCTATTTGCTATTAATAACTGCAGGGCGCCCTTCTCCTTG 1560
QY 1561 gcagcctgggtgttctccagtcggaagcctttagtgtatgataactacaactc 1620
Db 1561 GACAGCTGGGCTGTTCTCCAAGTGTCCGAGCCCTTGTAGTGTGATCAACTTACAACATC 1620
QY 1621 tctgcggctactcacgggtggccatctgtgaagtccaggattggatcttctgagatgctg 1680
Db 1621 TCTGCCGGCTACTCACGGGTGGCCATCTGTGAAGTCCCAGGATTTGGATCTTCGAGATGCTG 1680
QY 1681 tcttgcagaacaataactattacaacaaagaaacggagttgcaagctcaacatc 1740
Db 1681 TCTTTCAGAACAAATTTACACTATTACAAACAAAGAAAGCGAGTTGCAAAAGCTTCAACATC 1740
QY 1741 ttcttcaaggagctgaactataaaactaattcgagcttcccttctgtacagatggtcagc 1800
Db 1741 TTCTTCAAGAGCTGAACATATAAACTAATTCGGAGTCTCCCTTCTGTCTCAGATGGTTCAGC 1800
QY 1801 ctctgtccaactcctggcagcgagtgagcctgtggttttggctcgtccgtgctctctgtg 1860
Db 1801 CTTCCTGTCCAACCTGGCGACGAGTGGAGCCGTGTGTTTGGCTCGTCCGTCTCTCTGTG 1860
QY 1861 gtygagatggcggaagtcattcttcacactccttggttcatacacttctatgtgtacgc 1920
Db 1861 GTGGAGATGGCGACGCTATCTCGACCTCCCTGGTGTGATCATCACACTTCTCATGCTGTACGC 1920
QY 1921 cgggtccgagcgggtactggtctccaggaagggggccaggggtgccaagggaggtggcc 1980
Db 1921 CGGTTCCGAGCGGTACTTGGTCTCCAGACAGGGGGCCAGGGGTGCCAGGGAGGTGGCC 1980
QY 1981 tccactcgaagcttccctcccgctcccgcttctgtccctcaccctacatccccaccact 2040
|||||

Db 1981 TCCACTCCAGCTTCTCTCCGTCCCGTTCCTGTCTCCTCACCTACATCCCCACCACCT 2040
QY 2041 tctttcccccagagggcatgacccctccctggcctggcctgacagccctcacactgctat 2100
Db 2041 TCATTGGCCCCAGCAGGGCATGACCCCTGCCCTTGCCCTTGACAGCCCTCCACCTGGCTAT 2100
QY 2101 gctactctagggcccagtgccctccactggactctgctggcctgactgttctgctgt 2160
Db 2101 GCTACTCTAGGCCCCAGTGGCCCCCTCCACTGGACTCTCGCGGCGCTGACTGTCTGCGCTGT 2160
QY 2161 gccctggcgcgtctgagagagagaagatcctctcaaccagcctgagctccttgt 2220
Db 2161 GCCCTGGCGCGCTCTGAGAGAGAGAAGGATCCTCTACGCCAGCCCTGAGCTCCCTTGT 2220
QY 2221 aaactatgatatctcaacttcagcacatcttcccagcgcccccgagcagcctctgtg 2280
Db 2221 AAATGATGATATCTACTTTCAGCATCTTCCCCAGCGCCCCCGGAGCCCTCTGTG 2280
QY 2281 gttgctgctgagggaaacaggtctgtagtaaggggcccaagagcgtccagaggaagc 2340
Db 2281 GTGTGCTGCTGAGGGAACAGGCTGAGTAAGGGGCCAGGAAGCCGTCCAGAGGACAGG 2340
QY 2341 ggttaacgatctgctcagagcgtccctgccccctgtcttctgaacactgcttcccaagca 2400
Db 2341 GGCTAAGATCTGCTCAGAGCTGCCCTGCCCTGCTCTTCTGAACACTGCTTCCACAAGCA 2400
QY 2401 cagacaagttcccttccctcgatcagcccagcagacttggagctcagacaagaac 2460
Db 2401 CAGACAAGTTCCCTTTCCTTCGATCAGCCCGCCAGCACTTGGAGCTCTGACAAGGAAC 2460
QY 2461 ttctctgggagaggaaccaaggaataagacagccatcgactcctcaacccgtgacc 2520
Db 2461 TTTCTCTGGGAGAGGACCAGGAACAATAAGACAGCCATCGCACTCTCTACCCCGTGACC 2520
QY 2521 agaggtggcctggcctcactgcttcagggacactgagtctgctcctcctctgaact 2580
Db 2521 AGAGGCTGGCTGGCCTCACTGCTTTCAGGACACTGATGCTGCTCTCTCTTGAATC 2580
QY 2581 tgggtggggaaccccaccccaaaagccctttagttcttggcaattccccctccctc 2640
Db 2581 TGGGTGGGGAACCCCAACCCAAAGCCCTTGTAGTTCCTTGGCAATCCCCCTCCCTC 2640
QY 2641 actcctcaggtgggactagatgaagacagacaggttaggctggttctctctcat 2700
Db 2641 ACTCCTCAGGCTGGGACTAGAGTGAACAGACAGCAGGCTAGGCTGTCTCTCAT 2700
QY 2701 tctcaagactctctcatttgatacacatccccagtcgctctggtgtgtgtctcttc 2760
Db 2701 TCTCAAGACTCTCTCATTTGATACACCATACCCAGTGCCTCTGTGGTGTGTCTCCTTC 2760
QY 2761 ttctgttctcagcctggaagcttcttcaactgttccctcagaacaacagctctatgca 2820
Db 2761 TTTCTGTCTCTCAGCCTTGGAGCTTCTTTCAGTGTCCCTCAGAACACAGCTCTATGCA 2820
QY 2821 tcccttagaaacctgccaaaagcacattcttctgtgaactcagttacccctgtatgt 2880
Db 2821 TCCCTTAGAAACCTGCCAAAGGCACATTCTTTGTGTGAACTCAGTTACCCCTGTATGT 2880
QY 2881 ctccccagaattgtcgccctctccccctacccccacagtgactgtattatattgttccac 2940
Db 2881 CTCCCCAGAAATGTGCGCCCTCTCCCTCACCCCCACAGTGTACTGTATTATATGTTTAC 2940
QY 2941 attctttgtgtcgtcctcctgggttaaacctgagctccttctgtgtgctagggatgagatt 3000
Db 2941 ATTCTTTTGTGTGCTGCCCTGCCCTGGTAAACTGAGCTCTCTTGTGTGTGTCAGGATGAGATT 3000
QY 3001 tgcctgttttctgacttctccggtctagcccagcttcccacttgggagcgtgagcaggt 3060
Db 3001 TGTCTGTGTGTATCTCTCCGGTCTAGCCCACTTCCCACTTGGGACGGGTAGGAGGT 3060
QY 3061 actcaataaaggctgttctccatcaaaaaaataaaaaaataaaaaaataaaaaa 3117
Db 3061 ACTCAATAAAGGCTTGTTCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3117

Db	1362	CTCCATCAGTATGAGAAAGAACGCCCTGGCAGACCTCGGAGGAATATACGGCGACTGTAC	1321
Qy	1347	tgagaaaggtagcgatgctccgggtcaagaacottttaccotttcccaagtatacacagcaggt	1406
Db	1322	TGAGAAATGGTAGCGATGTCGGGTCAAGAACCTTTTACCCTTCCAAGTATACACAGCAGGT	1381
Qy	1407	gtgcattcactctctgtccaggagaaacatgatcaagaagtgtggctgtgctacatttt	1466
Db	1382	GTGCATTTCATCTCTGCTTCAGGAGAACATGATCAAGAAGTGTGGCTGTGCCTACATCTT	1441
Qy	1467	ctaccctaagcccaaggaggttgattctgtgtactaccgaagaagcagagcctctgggcta	1526
Db	1442	CTACCTTAAGCCCAAGGGAGTTGAGTTCGTGTACTTACCGAAGCAGAGCTCCTGGGCTA	1501
Qy	1527	ttgctattaataactcagggcgctctctcttgacagcctgggctgtttctccaaagt	1586
Db	1502	TTGCTATTATAAATCGCAGGGCCTTCTCCTTGGACAGCCTGGCTGTCTTCCAAGTG	1561
Qy	1587	tcggaagccttgtagtgtgatcaactacaaactctctgcggtactactoaogtggccatc	1646
Db	1562	TCGGAAGCCCTGTAGTGTGATCAACTTACAAATCTCTGCCGCTACTCACGGTGGCCATC	1621
Qy	1647	tgtgaagtccagattgtaattcttgagatgctgcttgcagaacaattacactattaa	1706
Db	1622	TGTGAAGTCCCAGATTTGATCTTCGAGATGCTGTCTTGCAGAACAAATTACATATTAA	1681
Qy	1707	caacaaagaaacggagttgcgaagctcaacatctctcaagagagctgaactataaac	1766
Db	1682	CACAAAGAAACGGAGTTGCAAGCTCAACATCTTCTTCAAGGAGTGAACATATAAAC	1741
Qy	1767	taattoggagctctctctgtcacagatggtcagcctctgttccaaacctgggcagccaagt	1826
Db	1742	TAATTGGAGTCTCTTGTACGATGGTCAGCCTCCTGTCCAACCTGGCAGCCAGTG	1801
Qy	1827	gagcctgtggttgctgcgtccgctctctgtgtgagatggcgagcgtcatcttcga	1886
Db	1802	GAGCTGTGGTTTGGCTGTCTCGTCTCTCTGTGTGGAGATGGCGAGCTCATCTTCGA	1861
Qy	1887	ctctctgtcatcacacttctcatgctgtagccgggttccggagccggtaactgtgtcc	1946
Db	1862	CCTCTGGTCAATCACATTCTCATGCTGCTACGCCGGTTCGGAGCCGTACTGTCTCC	1921
Qy	1947	agacagagggccaggggtgccaggagggtggctccactccactccagcttctctcccgtc	2006
Db	1922	AGGACAGGGGGCAGGGGTGCCAGGAGGTGGCTTCCACTTCAGCTTCTCTTCCCGTC	1981
Qy	2007	ccgttctgtctcacctcaactccccaccacttcttgcgccagagggcatgaccc	2066
Db	1982	CCGTTTCTGTCTTACCCTTACATCCCCACCACTTCTTTGGCCACAGGCGATGACCCC	2041
Qy	2067	tcacctggccctgacagccctccactgcctatgctactctaggccccagtgccccctc	2126
Db	2042	TCCCTTGGCCCTGACAGCCCTCCACCTTGCCTATGCTACTTAGGCCCCAGTGCCCCCTC	2101
Qy	2127	actggactctgcccgcctgaactgttctgctgtgccccgtggcgctctgagagagag	2186
Db	2102	ACTGGACTCTGCGGGCCCTGACTGTGTCTGCTGTGCCCTGTGCCGTCTGTGAGAGAGAG	2161
Qy	2187	aaggatcctctaccagccctgagctccctgttaaaactgattgatatctcaactcagea	2246
Db	2162	AAGGATCCTCTACCCAGAGCCCTGAGCTCCCTTGTAAACTGATGTGATATCTCATCTTACGA	2221
Qy	2247	catcttcccagcgccgcgcgagccctctgtgtgtgctgtgctgagggaaacaggtg	2306
Db	2222	CATCTTCCCAGCGGCCCGCGAGCCCTCTGTGTGTGTGCTTGTGCTGAGGGAACAGGCTG	2281
Qy	2307	agtaaggggccaggagccgtccagaggaaggggcttaacgatctctgcagagctgcc	2366
Db	2282	AGTAAGGGGCCAGGAAGCCGTCCAGAGACAGGGGCTAACGATCTGCTCAGAGCTGCC	2341
Qy	2367	tgcctctgtcttgaactgctcttc--cacaaagcaagaaagtctccctctcttcctcgg	2424
Db	2342	TGCCCTGTCTTGAACATGCTTCCACAAAGCACAGAACGTTGCCCTTCTTCCCTCGG	2401

[illegible]

Pennsylvania, 422 Curie Boulevard, 700 Clinical Research Building,
Philadelphia, PA 19104, USA

FEATURES

source
1. .3000
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="6"
/map="distal"
32. .2131

CDS

/codon_start=1
/product="epithelial sodium channel alpha subunit"
/protein_id="AAD21244.1"
/db_xref="GI:4469399"
/translation="MDHTRAPELNDLDVNSPKGMKNFKEDLCPPLPMQG
LGKDKREQLGPPSPPEEELIEFHSRYRELFOPFNNTHIGHAIRLVCS
KHNRMKFAWILNCTGMWQFALLEEYFVSYPLNINLNSDKDLVFAFVVTCL
NPYRTEIKARSDLEEDLRTTEOTLFOLYKNSYTRQAGRRSRDLGALPHPLQRL
RTPPPNPARSARSASSVRDNNPOVDKDKWIGFOLCNONKSDCFYOTYSSGVDAVR
EWRPHYINILSRPDTSPALEEALGSFIETCRENOAPCQANYSQFHHPMYGNCT
FNKNNSLNMSSMPGVNGLSLTURTQNDIFIPULSTVTGARVMVHGOEPAFMDG
GFNVRPGETSISMKEALDSLGGYNGDCTENGSDVPVKNLYPSKYTOQVCIHSCFQV
NMIRKCGAYIFYPKPGVEFDYLYKQSWGYCYKQLQAAFLSDLSLGFSKRPKPSV
TNYLKSAGYRWPSVKSDWIFEMLSLQNNYTIINNRNGVAKLNIFFRELNYKTSES
PSVTWVSLNSLGSOWSLWFGSSVLSVVEMAELIFDLIVITLIMLLHFRSRYWSPGR
GARGAREVASTPASSFPSPFCPHTSPSPSLPQOQTTPPLALTAPPPAYATLGPSPAS
LDSAVPGSSACAPANAL"

BASE COUNT 666 a 958 c 701 g 675 t

ORIGIN

Query Match 73.1%; Score 2280; DB 12; Length 3000;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 2674; Conservative 0; Mismatches 300; Indels 104; Gaps 11;

QY	51	cgagcagccccattcgtctcaacgctaataatgatgctggaccacacagagccccctgagct	110
DB	1	CGAGCGCCCCATTCTGCCTCCATACTAATGATGCTGGACCACACAGGCCCTCAGCT	60
QY	111	caacattgacctagacctcaacgctcccaactcgcctaagggtcctatgaagggaacaa	170
DB	61	CAACCTTGACCTAGACCTTGACCTGCCAACTCACCGAAGGACTCCATGAAGGGCAACAA	120
QY	171	attcaaggagcaagacctgtcctcctcagcccatgcaagactgggaagggaacaa	230
DB	121	TTTCAAGGAGCAGACCTTGTCTCTCTGCCCCATGCAAGGACTGGGCAAGGGGGACAA	180
QY	231	acgtgaaggagcaggcctggggccggaacccctcagcaccgcccgagccacccagagagga	290
DB	181	GGGTGAAGAACAGGCGCTGGGCGCGGAACCTTCAGAGCCCCGGCAGCCACGAGGAGGA	240
QY	291	ggaggcaactgattgaattccacgctcctacagggagctcttcagttcttcttgcaacaa	350
DB	241	GGAGGCACTGATCGAGTTCCACCGCTCTCCACCGGAGCTTCCAGTTCTTCTGCAACAA	300
QY	351	caccacatccacggggccatccgctggtgtcctcaaacacacacacacacacacacacgc	410
DB	301	TACCAACATCCACGGTGCATCCGCTGGTGTGCTCCAGACCAACCGCATGAAGACGGC	360
QY	411	cttctggcggtgctgtggctgtgaccccttggcatgactgactggcagcttcgcttgcct	470
DB	361	CTTCTGGCGGTGCTGTGGCTCTGCACCTTCGGCATGATGATCTGGCAGTTTCTTTGCT	420
QY	471	gttcggaggagtacctcagctaccagctcagcctcaacatcaacacacacacacacacac	530
DB	421	GTTTCGAGGAATACTTTAGCTTACCCCTGAGTCTCAACATCAACCTCAATTCGGCAAGCT	480
QY	531	ggtcttccctgcgctcactgtctgaccccttaactcctcaacacacacacacacacacac	590
DB	481	GGTCTTCCCTGCGGTGCTGTGTGTCACCCCTTAATCTCTTACAGATACACTGAAATTAAGA	540
QY	591	ggagctggaagagctggaccgcatcacggagcagacgcttttgcattgtacacacacacaa	650

DB	541	GGATCTGGAAGAGCTGGACCGCATACGGAAACAGACGCTTTTGAACCTGTACAAATACAA	600
QY	651	ctcttctacactcgcaggctggggcccgagcgcgagctcccgagacctctcgtgggtgc	710
DB	601	CTCTTCTTACTACTCGCAGGCTGGGGCGCGCCGACGACCCGCGACCTTCGGGGGTGC	660
QY	711	ttcccgacccccctcagcgctcgcactccactccgctcctactccgagcgacgcgc	770
DB	661	TTTCCACACACCCCTGCAGCGCTGCGACACCACTCCGCCCAATCCCGCCGCTCGGC	720
QY	771	gcgagcggtgtcttccagcgacgcgacacacacacacacacacacacacacacacacac	830
DB	721	GGCAGCGCGTCTTCCAGTGTACGCGACACAAATCCCAAGTGGACAGGAAGGACTGGAA	780
QY	831	gacgcttccaaactgcaacacagacacacacacacacacacacacacacacacacac	890
DB	781	AATCGGCTTCCAACTGTGAACACAGAACAAATCAGACTGCTTCTACGACATACATCAT	840
QY	891	tgggtggatgcagtgaggagtggtaccgcttccattccatcaacattctgtccagact	950
DB	841	CGGGGTGGATGCGCTGAGAGTGGTACCGCTTCCATTTACATCAACATTTCTGCCAGACT	900
QY	951	gtcggacacctcgcgcctcgtatagaggaagcgcctggggcaacttcatcttccacctgctg	1010
DB	901	GCCCGACACCTCGCTGCTCTAGAGGAAGACCCCTGGGCAGCTTCTATCTTTACCTGTCG	960
QY	1011	cttcaacacggccccctgcaacacgagcggaattatttccaaagtctccacacacacacac	1070
DB	961	TTTCAACACGCGCCCTGCAATCAGGCGAATATTTCAGTTCACACACCCCTGTATGG	1020
QY	1071	gaactgtacactttcaatgacaagaacacacacacacacacacacacacacacacacac	1130
DB	1021	GAACCTGTACACTTTCACACACAGAACAACTCCAATCTCTGATGCTTCTTCATGCTGG	1080
QY	1131	agtcacaatggtgtgtcctgacactgcgcacagagcagaatgacttccacctgct	1190
DB	1081	AGTCAACAATGTTGTCTCCCTGACACTGCGCACAGAGCAGAACTGACTTTCACCTCCCT	1140
QY	1191	gtccacagtgaacggggccagggtgatggtgcatggtcaggatgagcctccttatgga	1250
DB	1141	GTCACAGTGACGGGGGCCAGGTGATGTTGTCACGGTTCAGGATGAGCCTCTTTATGGA	1200
QY	1251	tgatggtgctcaacttgagcctggcgtggagacctccatcagtatgagaaagagc	1310
DB	1201	TGATGTTGCTTCAAGCTGAGGCTGGTGTGGAGACCTCCATCAGTATGAGAAAGGAGC	1260
QY	1311	cctggacagcctcgaggaaattacgcgcgactgtactgagaaatggtagcgtatcccggt	1370
DB	1261	CCTGGACAGGCTCGGAGGCAACTACGGAGACTGCACCTGAGAAATGGCAGCATGTCCCTGT	1320
QY	1371	caagaacctttacacctccaagtatacacagcagggtgtgcatcactcctgcttccagga	1430
DB	1321	CAAGAACCTTTACCCCTCCAAGTACACACAGCAGGTTGTCATTCCTCTCTCTCCAGGA	1380
QY	1431	gaacatgatacaagaagtgtgctgtgctcactctctcctcctcctcctcctcctcctcct	1490
DB	1381	GAACATGATCAAGAAGTGTGGCTGTGCCTACATCTTCTACCTTAAGCCCAAGGGGTAGA	1440
QY	1491	gttctgtgactaccgaagcagagctcctggggtatttgcattattataaactcagcggcgc	1550
DB	1441	GTTCTGTGACTACTTAAAGCAGAGCTCTCTGGGGCTACTGTACTATATAAATCGCAGGTGC	1500
QY	1551	cttctccttgacagcctgggtgttctccaaagtgcggaagcctgtagtgtgacaa	1610
DB	1501	CTTCTCTTGGATAGCTGGGTGCTTCTCCAAGTGCAGGAAGCCGCTGCGAGTGTGACCAA	1560
QY	1611	ctacaactctcgcgctcactcaggtggcctcctgtgaagtcccaggtatggactct	1670
DB	1561	CTACAGCTCTCTGCTGGCTTACTCAAGATGGCGCTCTGTGAAGTCCAGGATTTGGATCTT	1620
QY	1671	cgagatgctgtccttgacagaacattacactattacaacacacacacacacacacacacac	1730
DB	1621	CGAGATGCTATCTTGCAGAAATAATTACACGATCAACACAAACAAACGAGTTGCTAA	1680

FNDKNNLSNLMSPGVNNGLSLITLREQNEFIPLSTVTGARVWVHODEFAFMDG
GNLRPGVETISIMRKEALDSLGNVGDCTENGSDVPKNLYPSKYTOVCVTHSCFOE
NMKKGGCAVIFVPKPKGFEDYRKQSWGVCYKLOGAFSLDSLGCFSKRKPQSV
INYKLSAGYSRWPYSQDWIPEMLSLONNYINRKNVAKLINIFKELNKTNSES
PSVTWVLSLNSJGWSMLFSGSVLSVWEMAEILFDLLVITLLMLLRFRSRWSRGR
GARGAREVASTPASSFPSPHPTSPFPLSQQGMTPPLALTATPPPAYATLGPSAPP
LPSAAPDCSACALAAI"

BASE COUNT 480 a 699 c 545 g 476 t
ORIGIN

Query Match 70.5%; Score 2196.8; DB 12; Length 2200;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	47	gtcccgagacgccccattcgtctcaagctaatgatgtgagccacacagagccctg	106
Db	1	gtcccgagacgccccattcgtctcaagctaatgatgtgagccacacagagccctg	60
QY	107	agctcaacattgacctagaccttcaagcctccaaactcgctcaaggggtccatgaaggca	166
Db	61	agctcaacattgacctagaccttcaagcctccaaactcgctcaaggggtccatgaaggca	120
QY	167	accaattcaaggagcaagcccttgtctcctcagcccatgcaaggactgggaagggg	226
Db	121	accaattcaaggagcaagcccttgtctcctcagcccatgcaaggactgggaagggg	180
QY	227	acaaactgaagagcagggcgtggcccggaacccctcagcaaccccgagccacaggg	286
Db	181	acaaactgaagagcagggcgtggcccggaacccctcagcaaccccgagccacaggg	240
QY	287	aggagagcactgattgaattccaccgctcctaccggagctctccagttctctgca	346
Db	241	aggagagcactgattgaattccaccgctcctaccggagctctccagttctctgca	300
QY	347	acaaacacacattccacggggccatccgctggtgtgctccaaacacaaacgcgatga	406
Db	301	acaaacacacattccacggggccatccgctggtgtgctccaaacacaaacgcgatga	360
QY	407	cggcctctggcggtgctgtgctgtgaccccttgccatgagtgactgagctgcct	466
Db	361	cggcctctggcggtgctgtgctgtgaccccttgccatgagtgactgagctgcct	420
QY	467	tgtgttcgagagtagctcagctacccagtgagcctcaacatcaacctaactcagaca	526
Db	421	tgtgttcgagagtagctcagctacccagtgagcctcaacatcaacctaactcagaca	480
QY	527	agctgtcttcctgcctgacctgtctgaccccttaactcctacagatacactgaaatta	586
Db	481	agctgtcttcctgcctgacctgtctgaccccttaactcctacagatacactgaaatta	540
QY	587	aagagagctggaagagctggaacgcatcagcagagacgcttttgaactgtacaaat	646
Db	541	aagagagctggaagagctggaacgcatcagcagagacgcttttgaactgtacaaat	600
QY	647	acaactctctacactcgccagctggggccgagccgagcctcccgacactcctg	706
Db	601	acaactctctacactcgccagctggggccgagccgagcctcccgacactcctg	660
QY	707	gtgtcttcctgcgacccctgagcctgagcctccacccctcagcctcctccgagcga	766
Db	661	gtgtcttcctgcgacccctgagcctgagcctccacccctcagcctcctccgagcga	720
QY	767	cggcgagcgggttctccagcgtacgcgaacaaatcccccaagtggacgggaagact	826
Db	721	cggcgagcgggttctccagcgtacgcgaacaaatcccccaagtggacgggaagact	780
QY	827	ggaagatcgcttccaaactggaaccagaaacaaatcagactgtttctacagacatact	886
Db	781	ggaagatcgcttccaaactggaaccagaaacaaatcagactgtttctacagacatact	840
QY	887	ccctctgggtggatgtagtgaggagtggtaccgcttccattacatacaactctgtcca	946

Db	841	CTCTGGGGTGGATGCAGTGAGGGAGTGGTACCCTTCCATTACATCAACATTCTGTCCA	900
QY	947	gactgtcgggacacctcgcgcgtcttagaggaagaagccctgggcaacttcatctacact	1006
Db	901	GACTGTCCGACACCTGCCCGCTCTAGAGGAAGAAGCCCTGGGCACTTTCATCTCACCT	960
QY	1007	gtcgcttcaaccaggccccctgcaaccaggcggaattattccaagtctccaccacccccatgt	1066
Db	961	GTGCTTCAACAGGCCCTCCGCAACAGGCGGAATTATTCGAAGTTCACACCCCATGT	1020
QY	1067	acgggaactgtacactttcaatgacagaacaaactccaatctctggatgtctctccatgc	1126
Db	1021	ACGGAACTGCTACACTTTCAATGACAAGAACAACTCCAATCTCTGGATGTCTCTCCATGC	1080
QY	1127	ctggagtcacaatggtttgtccctgacactgcacagagcagaatgacttcatcccc	1186
Db	1081	CTGGAGTCAACAATGTTGTCCCTGACACTGCGCAGAGACAGAATGACTTCATCCCTC	1140
QY	1187	tgtgtccacagtgacgggggcccagggtgatgggtgcgtggtcaggatgagcctgccttta	1246
Db	1141	TGCTGTCCACAGTGACGGGGGCCAGGGTGATGGTGCATGGTCAGGATGAGCCTGCCCTTA	1200
QY	1247	tgatgtgtgtggttcaacttgggctggcgctggagacctccatcagtatgagaagg	1306
Db	1201	TGGATGATGGTGGCTTCAACTTGAGGCTGCGGTGGAGACCTCCATCAGTATGAGAAAGG	1260
QY	1307	aagccctggagcagcctcggaggaattacggcgaactgtactgagaatggtagcatgcc	1366
Db	1261	AAGCCCTTGGAGAGCCTCGGAGGAAATTTACGGCGACTGTACTGAGAATGGTAGCATGTCC	1320
QY	1367	cgttcaagaacaccttacccttccaaagtatacacagcaggtgtgcatcactcctgcctcc	1426
Db	1321	CGGTCAAGAACCTTTACCTTCCAAAGTATACACAGCAGGTGTGCATTCACCTCTGCTTCC	1380
QY	1427	aggagaacatgatcaagaagtgtggctgctcactatcttccacccaagcccaaggag	1486
Db	1381	AGGAGAACATGATCAAGAAAGTGTGGCTGTGCCTACATCTTACCTTAAAGCCCAAGGAG	1440
QY	1487	ttgaattctgtgactaccgaaagcagagcctcctgggctattgctattataaaactgagg	1546
Db	1441	TTGAGTTCTGTGACTACCGAAAGCAGAGCTCCTGGGCTATTGTCTATTATAAATCGCAGG	1500
QY	1547	gcgcctctcctgtgacagcctgggctgttctccaaagtgtcgaagccttgtagtgga	1606
Db	1501	GGCCTTCTCCTTGGACAGCCTGGGCTGTTTCTCCAAAGTGTCCGAAGCCTTGTAGTGTGA	1560
QY	1607	tcaactcaaacactctcgcggctactcaoggtggccactctgtgaagtcaccagatgga	1666
Db	1561	TCAACTACAAACTCTCTGCCGGCTACTCACGGTGGCCATCTGTGAAGTCCCGAGATTGGA	1620
QY	1667	tcttcgagatgctgtccttgcagaaacaaattacactattaacaaacaaagaaacaggattg	1726
Db	1621	TCATTTCGAGATGCTGCTCTTGCAGAACAAATTACACTATTAAACAAACAAAGAGAGTTG	1680
QY	1727	caagctcaacatcttctcaaggagctgaactataaaactaatcggagctcctctctt	1786
Db	1681	CAAAGCTCAACATCTTCTTCAAGGAGCTGAACATATAAACTAATTCGAGCTCTCTTCTG	1740
QY	1787	tcacgatggttagcctcctgtccaaacctgggagccagtcagtcgtgtgtgtgtgtgtgt	1846
Db	1741	TCACGATGGTCAGCCTCTCTGTCCAACCTGGGCGCAGCCAGTCGAGCTGTGGTGTGGCTCGT	1800
QY	1847	ccgtgctctctgtgtgtgagatgagcagctcatcttcgaacctcctggctacacacttc	1906
Db	1801	CCGTGCTCTCTGTGGTGGAGATGGCGAGCTCATCTTCGACCTCTCTGCTCATCACACTTC	1860
QY	1907	tcatgctgtacgcgggttccggagcggtagctgtctcagagcagggggccagggtg	1966
Db	1861	TCATGCTGTACGCCGGTTCGGAGCCCGTACTGCTCTCCAGAGACGAGGGGCCAGGGGTG	1920
QY	1967	ccaggagggtggcctccactcagcttctccttcccgctcccgcttctctgtcctcaccccta	2026
Db	1921	CCAGGGAGGTGGCCTCCACTTCCAGCTTCTCTTCCCGTCCCGTTTCTCTCTCTCACCCCTA	1980

QY 2027 catccccacacactcttcttggccccagcaggcgatgaccctccctggcctgacagccc 2086
|||||
Db 1981 CATCCCCACACCTTCTTTGGCCCCAGCAGCGCATGACCCCTCCCTGGCCCTGACAGCCC 2040
QY 2087 ctccacactgcctatgctactctagccccagtcgccccctccactgcactgagcgccctg 2146
|||||
Db 2041 CTCCACCTGCCCTATGCTACTCTAGGCCCCAGTGCCTCCACTGGACTCTGGCGGCCCTG 2100
QY 2147 actgtctacactgtgccccgagcgctctgagagagagagagagatccctccaccagcc 2206
|||||
Db 2101 ACTGTTCTGCCCTGGCCCTGGCGGCTCTGAGAGAGGAGAGGATCCTCTCACCAGGCC 2160
QY 2207 ctgagctcccttgtaaaactgagtgatctcactctcagca 2246
|||||
Db 2161 CTGAGCTCCCTGTAAAGTATGGATATCTCATTCTCAGCA 2200

RESULT 5
RNU54699 2200 bp mRNA ROD 14-JUN-1997
LOCUS Rattus norvegicus epithelial sodium channel alpha subunit (rEnaca)
DEFINITION mRNA, complete cds.
ACCESSION U54699
VERSION U54699.1 GI:2148925
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2200)
AUTHORS Kreutz,R., Struk,B., Rubattu,S., Hubner,N., Szpirer,J., Szpirer,C.,
Ganten,D. and Lindpaintner,K.
TITLE Role of the alpha-, beta-, and gamma-subunits of epithelial sodium
channel in a model of polygenic hypertension
JOURNAL Hypertension 29 (1), 131-136 (1997)
MEDLINE 97191134
REFERENCE 2 (bases 1 to 2200)
AUTHORS Kreutz,R.
TITLE Direct Submission
JOURNAL Submitted (11-Apr-1996) Reinhold Kreutz, Medicine, Medicine,
Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115,
USA

FEATURES
source
1..2200 Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Wistar-Kyoto, Heidelberg (WKY-1HD)"
/db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="kidney"
/chromosome="4"
36..2132
/gene="rEnaca"
36..2132
/codon_start=1
/product="epithelial sodium channel alpha subunit"
/protein_id="AA061156.1"
/db_xref="GI:2148926"

gene
CDS

Query Match 70.4%; Score 2195.2; DB 12; Length 2200;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2197; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 47 gtccccgacagcccccttctgcttcaagctaatgatgctggagccacacacagagccctg 105
|||||
Db 1 GTCCCGACAGCCCCATTCGCTTCCAGCTAATGATGCTGGACACACACAGAGCCCCCTG 60
QY 107 agctcaacattgacctagaccttcacgcctcccaactcgcctaaaggggtccatgaaggcca 166
|||||
Db 61 AGCTCAACATTGACCTTAGACCTTACGCCCTCCAACTCGCTAAGGGGTCCATGAAGGGCA 120
QY 167 accaattcaaggagagagacccttgcctcctcagccccatgcaagacatggggaagggg 226
|||||
Db 121 ACCAATTCAAGGAGCAAGACCCCTTGCTCTCCTCAGCCCCATGCAAGGACTGGGGAAGGGG 180
QY 227 acaaacgtgaagagcagggcctgggcccgaacccctcagcacccccggcagccaccagag 286
|||||
Db 181 ACAAACTGAAGAGCAGGGCCTGGGCCGGGAACCTCTCAGACCCCGGCACCCACCGAGG 240
QY 287 aggagggagcactgattgaattccaccgcctcctaccgggagctcttccagtcttcttgc 346
|||||
Db 241 AGGAGGAGGCACCTGATTGAGTTCACCGCTCTTACCGGGAGCTCTTCCAGTCTTCTGCA 300
QY 347 acaaacaccaccatccacggggccatccgcctgggtgctccaaacacacacccgcatgaaga 406
|||||
Db 301 ACAAACTGAAGAGCAGGGCCTGGGCCGGGAACCTCTCAGACCCCGGCACCCACCGATGAAGA 360
QY 407 cgagcctctggcggtgctgctgctgctgacacctcgccatgagctgagctgagctgcacct 466
|||||
Db 361 CGGCCCTCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 467 tgcgttcgagagtagtacctcagctaccagctcagctcagctcagctcagctcagctcagctc 526
|||||
Db 421 TGCTGTTGAGGAGTAGTACCTCAGCTACCCAGTGAGCTCAACATCAACCTCAATTACAGACA 480
QY 527 agctgtcttccctgcgctcactgtgtgaccccttaactccttacagatacacatgaaatta 586
|||||
Db 481 AGTGTGCTCTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 587 aagagagctggaagagctggaccgcacatcacgagcagcagcagctttttgacttatacaaa 646
|||||
Db 541 AAGAGAGAGCTGGAAGAGCTGGACCGCATCACGGAGGAGAGCCCTTTTGTACTTGTACAAAT 600
QY 647 acaactcttctctacactgcagcagctggggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 706
|||||
Db 601 ACAACTCTTCTTACTACCTGCGCAGGCTGGGGCCCGCAGCGCGCAGCTCCCGGCGACTCTCTGG 660
QY 707 gtgcttcccgacccccctgcagcgcctgcgcgcctccacactcgcgcctactcgcgcgcgcgc 766
|||||
Db 661 GTGCTTTCCCGCACCCCTGTCAGCGCTGCGCCTGCGCCTGCGCCTGCGCCTGCGCCTGCGC 720
QY 767 cggcgagcagcggtcttccagcgtacgcagcaacaatcccgaagtggaccggaagagact 826
|||||
Db 721 CGCGCGCAGCGGCTCTTCCAGCGTACGGGACAACTCCCGGAGTGAGCCGCGGAGGACT 780
QY 827 ggaagatcggtctcccaactgtgcaacacagacaaaatcagactgttctaccagacataact 886
|||||
Db 781 GGAAGATCGCTTCCAACTGTGCAACAGACAGAAATACAGACTGTTTCTACACACATACT 840
QY 887 cctctgggtgagtgagtgagggagtggtaccgcttccattacatacaattctctgcca 946
|||||
Db 841 CCTCTGGGTGGATGAGTGAGGAGTGGTGTACCGCTTCCATTACATCAACATCTTGTCCA 900
QY 947 gactgctggagacactgcgcgcctctagaggaagacccctgggcaacttcaattctcaact 1006
|||||
Db 901 GACTGTGCGACACTCGCCGCTCTTAGAGGAAGAGCCCTGGGCAACTTTCATCTTCACT 960
QY 1007 gtcgcttcaaccagggccccctgcacacagcggaattattccaagttccaccaccccatgt 1066
|||||

BASE COUNT 479 a 699 c 546 g 476 t

Db	961	GTCCGCTTCAACAGGCCCCCTGCAACCGAGCGGAATTAATTCCAAGTTCACACACCCCAATGT	1020
Qy	1067	acggaaactgtacactttcaatgaacaagaacaactcaactctctgtagtgcctcccatgc	1126
Db	1021	ACGGGAAGCTGCTACACTTTCATATGACAAGAACAACTCCAATCTCTGGATGTCTCCATGCG	1080
Qy	1127	ctggagtcacaataggtttgtccctgacactgcgcacagagcagaatgaacttcaccccc	1186
Db	1081	CTGGAGTCAACAATGGTTGTCCCTGCACACTGCGCACAGACAGATGACTTCATCCGCC	1140
Qy	1187	tgctgtccacagtgacggggggccagggtgatgtcatgttggtcaagatgagcctgccttta	1246
Db	1141	TGCTGCCACAGTAGCGGGGGCCAGGGTGATGGTGCACTGGTCAAGATGAGCCTGCCCTTTA	1200
Qy	1247	tgatgatggtggcttcaacttgagcctggcgtagagacctccatcagtatgagaaagg	1306
Db	1201	TGGATGATGTTGGCTTCAACTTGAAGCCTGGCGTGGAGACCTCCATCAGTATGAGAAAGG	1260
Qy	1307	aagccctggacagcctcggaggaattacggcgactgtactgagaatggtagcagtgcc	1366
Db	1261	AAGCCCTGGACGCTCGGAGGAAATACGGCGGACTGTACTGAGAATGGTAGCATGTCC	1320
Qy	1367	cggtcaagaacctttacccttccaagtatacacagcaggtgtgaatcaactcctgttcc	1426
Db	1321	CGGTCAAGAACCTTTACCTTCCAAAGTATACACAGCAGGTGTGCATTCACTTCGTGCTTC	1380
Qy	1427	aggagaatgatcaagaagtgtggctgtgcctacatcttcaccctaagccaaaggagag	1486
Db	1381	AGGAGAACATGATCAGAAGTGTGGCTGTGGCTACATCTTACCCCTAAGCCCAAGGGAG	1440
Qy	1487	ttgagttctgtactaccgaagcagctcctgggctattgctattataaacctgcagg	1546
Db	1441	TTGAGTTCTGTGACTACGAAAGCAGAGCTCCTGGGCTATTGCTATTATTAATGTCAGG	1500
Qy	1547	gcgcttctccttggaagcctggggtgttctccaagtgtcgaagcctgtagtgtga	1606
Db	1501	CGCGCTTCTCCTTTGGACAGCCTGGGCTGTTCTCCAAGTGTGGAAGCCTTGTAGTGTGA	1560
Qy	1607	tcaactcaaatctctgcggctactcacgtggccatctgtgaagtcctccagatggga	1666
Db	1561	TCAACTACAACCTCTCTGCGGGCTACTACGGTGGCCATCTGTGAAGTCCCAGAGATTGGA	1620
Qy	1667	tcttcgagatgctgtccttgagacaattatacattatacaacaagaagaacgcaggttg	1726
Db	1621	TCTTCGAGATGCTGTCTTGCAACAATACACTATTAAACAACAAAGAAACGAGTTG	1680
Qy	1727	caagctcaacatctcttccaaggagctgaactataaacataattcggagctctcctctg	1786
Db	1681	CAAAAGCTCAACATCTTCTCAAGAGCTGAACATAAATAACTTAATTCGGAGTCTCCTCTTG	1740
Qy	1787	tcagatggttcagcctcctgtccaactgggcagcgatggaagcctgtggttgagctcgt	1846
Db	1741	TCAGATGTHCAGCCTCTGTCCAACCTGGGAGCCAGTGAGGCTGTGGTTGGCTCGT	1800
Qy	1847	ccgtgctctctgtgtggagatggcggacgtcatcttcgacctcctgtgtcatcaccttc	1906
Db	1801	CCGTGCTCTCTGTGGAGATGSGGAGCTCATCTTCGACCTCTCTGTCTATCATACTTC	1860
Qy	1907	tcatgtctacgccggttccggagccggtagtctgtgtccagcagcaggggccaggggtg	1966
Db	1861	TCATGCTGTCTACGCCGGTTCGGGAGCCGCTACTGTGTCTCCAGGACAGGGGGGTG	1920
Qy	1967	ccaggaggtggcctcacctccagcttctccttcccggttcgcttctgctcaccta	2026
Db	1921	CCAGGGAGTGGGCTCCACTTCCAGCTTCTCTCTCTCCCGTCCGTTTCTGTCTCTCACCTTA	1980
Qy	2027	catcccaaccaccttctttgccccagcaggcatgacccctcccctggcctgacagccc	2086
Db	1981	CATCCCCACCACTTCTTTGCCCCACGAGGGATGACCCCTCCCTTGCCCTTGACAGCCC	2040
Qy	2087	ctcacctgcctatgtactcttagggcccaagtgcctctccactggactctgtcgcgctg	2146
Db	2041	CTCCACTGTGCTATGTACTCTAGGCCCCAGTGCCTCTCCACTGTGACTCTGGCGGCTG	2100

Qy 2147 actgtttctgccttgcctcggcgccgctctctagagagagagaagtacatctctcaaccgcc 2206
Db 2101 ACTGTTCTGCCTGTGCCCTTGGCGGCTCTAGAGAGAGAAGGATCCTCTCACCCAGCC 2160

Qy 2207 ctgaactcccttgaaactgatgatatttcaccticagca 2246
Db 2161 CTGAGTCCTTGTAACATGATGGATACTCACTTCAGCA 2200

RESULT 6
HSLASNA 3151 bp mRNA PRI 01-FEB-1994
LOCUS H.sapiens mRNA for lung amiloride sensitive Na+ channel protein.
DEFINITION X76180
VERSION X76180.1 GI:452649
KEYWORDS Na+ channel; Na+ channel protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 3151)
Barbry,P.
Direct Submission
Submitted (19-NOV-1993) P. Barbry, CNRS, IPMC 660 Route des
Lucioles, 06560 Sophia Antipolis, FRANCE
2 (bases 1 to 3151)
Volley,N., Lingueche,E., Champigny,G., Mattei,M.G., Waldmann,R.,
Lazdunski,M. and Barbry,P.
The lung amiloride-sensitive Na+ channel: biophysical properties,
pharmacology, ontogenesis, and molecular cloning
Proc. Natl. Acad. Sci. U.S.A. 91 (1), 247-251 (1994)
94105144

Location/Qualifiers
1 . 3151
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/tissue_type="lung"
/cell_type="epithelial"
/dev_stage="adult"
100 . 2109
/codon_start=1
/product="Na+ channel protein"
/protein_id="CAA53773.1"
/db_xref="GI:452650"
/db_xref="SWISS-PROT:P37088"
/translation="MEGNKLEEDQSPPQTGPKLMKNKEEGCLGPPEAAPQPQAEF
EEALIEFHRSYRELFPFCNNTTHGAIRLSOHNRMKTAFAWILMCLTFGMGYMFQ
GLRGEYSVPVLNI NLSNDKLVPAAVTICTLNLPYPETKELELDRTTEQLD
LYKSSFTLVAGSRSDRLRGTLTPSHGVDAVEWRPHYINLLSPETFLSLEDITLG
WKMDQIFQLCNKNSDCFYQTSYSGVDVAWEVRIYINLLSPETFLSLEDITLG
NFIPACFNQVCNQARYSHFHPHMVNCYTENDNNLSNMMSMPGINGLSLMLRA
EQNDIPLLSTVGARVMHQDEPAFMDDGFNLPGVETSISMRKETIDRLGGDYGR
DKTKNGSDPVENLYPSKYTQOVCIHSCFOESMIKCGCAVIFYPRQNVEYCDQRKH
SSMGCYVKIQVDPSDDHCKFKPCSVTSVQLSAGYSRMPSVTSQEWVFQMLSR
QNNTYANKRNQAKVNI FKELNKYNINSPSVTVWTLNLSLGSQSNLFGSSVLSYLR
VEMAEIVDLLLIIMFLMLRRFRSYWSPGRGGGAQEAVASTLIASSPSPHFCHPMPLSI
SLSQPGAPPAPLATPAYATAYATGPRSPGGSGASSSTCTPLGGP"

BASE COUNT 677 a 995 c 803 g 676 t
ORIGIN

Query Match 49.2%; Score 1532.2; DB 9; Length 3151;
Best Local Similarity 73.3%; Pred. No. 0;
Matches 2254; Conservative 0; Mismatches 678; Indels 145; Gaps 17;

Qy 151 gggtccatgaaggccaaccaatcaagacagcaaaccttgtctctccacgccatgcaa 210
Db 94 GTCTCATGGAGGGAACAAGCTGGAGGACGAGACTTAGCCCTCCACAGTCCACTCCA 153

Qy 211 ggaactgggaagggggacaacagtgaagcagggcctggggccgggaacctcagcaccc 270
Db 211 ggaactgggaagggggacaacagtgaagcagggcctggggccgggaacctcagcaccc 270

GCTTCKRKPQSVITYIKLSASYSQWPSATSDQWTFQMLSRONNYTIKKRDRGVAKINTL FKELUNYSNSESPTVMTLLSNIGSQWLSWFGSSVLVSVEAMELIIDLLVITFLMLLL RRFRSRYSPGRGKGTQEVASTPAALSPSPFPHPAFTSSSPDPAISPALSPAPPA YATLGHPPAPSGLAEAATSAHAGEP*									
BASE COUNT	653 a	1239 c	896 g	798 t					
ORIGIN									
Query Match	45.0%;	Score 1401.8;	DB 3;	Length 3586;					
Best Local Similarity	72.4%;	Pred. No. 0;							
Matches 2035;	Conservative	0;	Mismatches	657;	Indels	119;	Gaps	12;	
QY	199	caGCCcAtgcaaggaCtgggaaggggacaaacagtgaaagagcggggCctggggcccgaaa	258						
DB	691	CTGACcTGTcAGGGCTcATgAAGGAGAcAAGcCTGAGGAGCGGGGCGGGTCCCGAG	750						
QY	259	ccctcagcaccgccgcagccacagagagagagagcagcattgaattccacgcctcc	318						
DB	751	CCTTCGGGCCCCCGCCCGCCAGGAGGAGGAGGCGCTGCTCGAGTTCACCGCTCC	810						
QY	319	tacgggagctcttcacgttcttctgcaacacaccaccatccacggggccatccgcctg	378						
DB	811	TACCGGGAGCTCTCGAGTCTTCTGCAATACACcACCATCCAGGGCGCCATCGCGCTG	870						
QY	379	gtgtctccaaacacacacgcgcagtgaaagcgccctctggcggtgctgtgctgtgcacc	438						
DB	871	GTGTCTCGCAGCACAATCGCATGAAGACAGTCTTCTGGGCGGTGCTGTGCTCTGCACC	930						
QY	439	ttcggcatgatactggcagttgcgctgtgttcggaggagtaacctcagctaccacgtg	498						
DB	931	TTTCGGCATGATGTACTTGGCAGCTTCGGCCAGCTTCTCGGAGAGTACTTCAGCTACCCCGTC	990						
QY	499	agctcaacataacctaattccagacaagctggtcttcctccctgcctgcactgtctgcacc	558						
DB	991	AGCTCAACATCAACCTCAACTCGGACAAGCTTGTCCTCCCTGCGGCTCTCCATCTGCACC	1050						
QY	559	cttaactcttcagatacactgaaattaaaaggagagctggaagagctggacgcgcatcacg	618						
DB	1051	CTAAACCCCTACAGGTACAAAGAAATTCAGAGGAGCTGGAGGAGCTGGACCGCATCAGC	1110						
QY	619	gagcagacgcttttgacttgtacaaaatacaactcttcactacactgcagagctggggcc	678						
DB	1111	GAGCAGACTCTTTTCGACCTGTACAAGTACAACTCCTCCAAGACCTC-----GTGGCC	1164						
QY	679	gcagcgccgagctcccgacacctcggtgctttccgcgcaacctctcagcgctcgccgcg	738						
DB	1165	CAGCGCCGCTCCCGCGGACCTGGGAGCGGCTGCGCGCACCCCTTCAGCGCTCGCC	1224						
QY	739	actccactccgcccctaactccggcgacgcgcgcagcgagcggtttccagcgtacgcgac	798						
DB	1225	GTCCCGCTCCGCCACGCGCGCGGAGTCCGCGCGCGCGGTCCAGCATGGCGGAC	1284						
QY	799	acaaatccccaaagtggacggaagactggaaagatcggtcttcactgtcaaccagaaac	858						
DB	1285	AACAACCCCAAGTGAACAGGAAGACTTGAAGATCGGCTTCCAGCTGTGAACCCAGAAC	1344						
QY	859	aaatcagactgtttaccagacataactcctctggggtggatgagtagagagtggtac	918						
DB	1345	AAATCGGACTGCTTCTACACAGCTACTTCGACGGGGTGGATGCTGTGAGAGAGTGGTAC	1404						
QY	919	cgcttccattacatacaactctgtccagactgtc--ggacacctgcgcgcctctagag	975						
DB	1405	CGTTCCACTACATCAACATCTGTCCAGCGGAGACAGGACACCTCTCCCTCGCTGGAG	1464						
QY	976	gaagaagccctgggcaacttcacttcacctgtgccttcaaccagggccccctgcaaccag	1035						
DB	1465	GAGGACGTGCTGGGCAAAATTCATCTACCTGCGCTTCAACACGAGACTCCTGCAATGAG	1524						
QY	1036	gcgaattattcgaagtccaccaccccatgtacgggaactgtcaacttcaatgacaag	1095						
DB	1525	GCGAATTACTCTCAATTTCCACCACCGGATGATGGAAACTGTACACTTTCAACGACAA	1584						
QY	1096	aacaactccaactctctgtagtgcctccatgcctggagtcaacaaatgggttgccttgaca	1155						

[illegible]

Db	1426	TGCTGTTACTCGCGATGGCCCTCGGTGACGCTCCAGGACTGGGTCTTCCAGATGCTGTC	1485
Qy	1683	cttgagacaattacactattataacaacaaagaaacaggagttgcaagactcaacatctt	1742
Db	1486	GCTACAGAACACTACACGGGTACGCAACAAGAGAATGGGTGCAAGCTCAACATCTA	1545
Qy	1743	cttcaaggagctgaactataaactaaatcggagctctcttctgtcacgagtggtcagcct	1802
Db	1546	CTTCAAGAGTTGAACACTACAAGCCAAATCTGAGTCCCTCTGTTCACGATGGTCACCT	1605
Qy	1803	cctgtccaacctgggcagcagtgagcgtgtggttttgctgtcctgtcgtctctgtgtgt	1862
Db	1606	CCTGTCCAACCTGGGCAGCAGCTGAGGCTGTGGTTCGGCTCTCCGTCTGCTCGGTGT	1665
Qy	1863	ggagatggcggagcctcaattcttggaacctcctgtgtaatacacactctcactgctgcgcgc	1922
Db	1666	GGAGATGGCCGAGCTCCTTCGACCTCTCAGTCATCATCTTCATGCTGCTCCGGCG	1725
Qy	1923	gttcggagcgggtactggtctccaggagcagggccaggggtgccaggaggtggcctc	1982
Db	1726	GTTCCGCGCCGCTACTGCTCTCTGGCGGGGGCGCGGGGGCGGGAGGTGGCCTC	1785
Qy	1983	cactcagcttctcctctcccgctccggtttctgtctcactacactatataccccacccttc	2042
Db	1786	CTCGCGGGTCTCCGGCTGCCCTCCCGTCTCTGCCCCACCCACGCTCC-----CCGTC	1839
Qy	2043	tttgccccagcagggcatgacccctccccctggcctgacagccccctcaactgcctatgc	2102
Db	1840	CGTGGCCCGAGCGTGCCGCCACACTGCCCCCGCTCTTGACAGCCCTCCACCTGCCATGC	1899
Qy	2103	tacttagccccagtgccccctccactgactcgtcggcgcctgactgcttctgctctgtgc	2162
Db	1900	CACTTTCGGCCCTGCCGTCTCAGTCCGGC-----TCGCTGTGTGC	1941
Qy	2163	cctgcygcygctctgagagaggaagatcctctcaccagcctgagctgcctctgttaa	2222
Db	1942	CCTGGGGAGCCCTAAGAGGAGACAGTGCCTCCCGCCCGCCCGAGCGGTG----CCAGGGGAA	1997
Qy	2223	actgatgatatctcacttcagacacatcttcccagcggccccgcggagcctctctgtgt	2282
Db	1998	GCCTTGGCAGAATGAAGATGTTTGGGGCTTTCCTCTCAGAGTGGCCCACTGCCCATGGT	2057
Qy	2283	gtgcctgctgagggaaacagcgtgtagtaaggggccaggaagccgtccagagagcagggg	2342
Db	2098	GGGC-----GGGGAAGGGGCCCCCGGAAGTTGTCAGAGAGCAGCGG	2098
Qy	2343	ctaacga-tctgctcagagctgcccctgcccctgtcttgaacactgcccctccacaagcac	2401
Db	2099	CCGGTGAGGCTGCCAGAACTGCCCTGTGCTGTGCCGGTTCCTCTGTGCTCC--CCTG	2156
Qy	2402	agacaagttccctctcctcctggatcagcccaagccagacttgagctctgacaaagaaact	2461
Db	2157	AACACTGGCGTCCCTCCCAAGTGCAGCGCCGAGCCAGACTGGGAATGCTGACAAGGAAC	2216
Qy	2462	ttccttggagagaccagaag----aacaataagacagcctcgcactcctaccctccgtg	2517
Db	2217	TTTCTAGGAGACGACCAACAGCCCAACACAAACAAAGAGCCCAAGAGGCGTGCACAGGTC	2276
Qy	2518	acc-----agaggtggcctggcctcactgcctcactgccttcaaggacactgat	2559
Db	2277	TCCTACGCAATGGTGACTGATGCCAGAGCCAGCCCGCCGCGCTGGCTGGGGCCACAGAT	2336
Qy	2560	gtcgtcctcctccttgaaacttggttggggaacccccccccccccccccc-ttgttagtt	2618
Db	2337	GTCCTGCTGTGCTTGAACCTGAGTGGGGAACCCCAACCCCAAGGCCCTTATTTCGTT	2396
Qy	2619	ctttggattccccctccccctcactcctcaggttggtgggactaggttaagacagcaggg	2678
Db	2397	CTTGGCAATTCCTCTCCCTGGCTGCGCAGGGTGAAGGCTGGAGCAGAGCGCTGTGTGTA	2456
Qy	2679	taggctgg 2687	
Db	2457	CAGGCGGG 2465	

RESULT 11

AF071230

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 2099;

Qy 109

Db 495

Qy 169

Db 555

Qy 229

Db 615

ctcaagcattgacctagaccttcacgcctcaactgcctcaagggctccatgaaggcaac

ctcaccgacggcgcccccctctgctgctgcacacctgcacggcctcaatgaagggtgat

caattcaaggagcaagacccttgctcctcctcagccctgacggactgagggaggggac

GAGCTCAAGGCACAGGCCCTCTCCTCCAGCCCTCCAGGGCCCTCAAGGGGGAC

aaactgagagcagggccttggtggcccggaacccctcagcccccggagccccaccggag

AAGTGTGACAGCCAGGGCTGGGACCTGAGCCCAACCGCCCGCCAGCAGCAGGAGG

43.5%; Score 1357; DB 12; Length 3385;

69.3%; Pred. No. 0;

5; Mismatches 760; Indels 163; Gaps 13;

109

495

169

555

229

615

ctcaagcattgacctagaccttcacgcctcaactgcctcaagggctccatgaaggcaac

ctcaccgacggcgcccccctctgctgctgcacacctgcacggcctcaatgaagggtgat

caattcaaggagcaagacccttgctcctcctcagccctgacggactgagggaggggac

GAGCTCAAGGCACAGGCCCTCTCCTCCAGCCCTCCAGGGCCCTCAAGGGGGAC

aaactgagagcagggccttggtggcccggaacccctcagcccccggagccccaccggag

AAGTGTGACAGCCAGGGCTGGGACCTGAGCCCAACCGCCCGCCAGCAGCAGGAGG

Db 736 CAGACGTACTCTCAGGGGTGGATGCGGGAATGTTACAGCTTTCACTACATCAAT 795
Qy 937 attctgtccagactgtcgacacctcgccgctctagagaaagaacccctgggcaacttc 996
Db 796 ATCTGGGACAGATGCTGATGCC---AAAGACCTGGATGAGTCTGACTTTGAGAAATTT 852
Qy 997 atcttcacactgtcgtctcaacaggcccccctgcaacaggcggaattattccaagtctccac 1056
Db 853 ATTATATGCTTGGCGCTTCAATGAAGCAACATGTGACAGGGAATTACACCCACTTCCAC 912
Qy 1057 caccacctataggggaactgtacactttcaatgacaagaacaactccaatctctggatg 1116
Db 913 CATCCCTTGTATGGGAACGTCTACACTTCAATGACAATAGCAGC---AGCCTGTGGACA 969
Qy 1117 tcttcacactgtggagtcacaatggtttgtccctgacactgcgcacagcagaatgac 1176
Db 970 TCTCTCTGCCGTGGGATCAATATGGCCCTCTCTCGTGGTGGCCACCGCAGCAGATGAT 1029
Qy 1177 ttcatcccccctgtgtccacagtgcagggggccagggtgatggtcatggtcaggatgag 1236
Db 1030 TTATCCCTCTGTGTCACGGTAACTGGAGCCAGGGTCATGCTCCACGATCAGATGAG 1089
Qy 1237 cctgctttatgagatggttggtctcaacttgagcctggcgtggagacctccatcagt 1296
Db 1090 CCAGCCTTTCATGGGCTTCAATGTGCTGCCAGGTATTTGAGACCTCCATCAGC 1149
Qy 1297 atgagaagaagccctgacagcctcgagagaaataacgggagctgtactgagaatggt 1356
Db 1150 ATGAGAAGGAGATGACTGAGCGTCTTGGGGGAGTTACAGCGACTGTACAGAGGATGGC 1209
Qy 1357 agcgatgtcccgttcagaacacctttacccttccaagtatacacagcaggtgtgtcattcac 1416
Db 1210 AGCGATGTCCAGTGCAAAATCTGTACTCATCTCGTACATCAACAGGTCTGCATTCGT 1269
Qy 1417 tectgtctccaggagaacatgatcaagaagtgtggtgtgcctacatactttaccctaaag 1476
Db 1270 TCTGTCTTTCAGCTCAACATGTTAAAGCGCTGCTCTGTGCTTATTTATCCCTTA 1329
Qy 1477 cccaaggagattgagttctgactaccgaagaagcagagctcctgggctatgctattat 1536
Db 1330 CTGTATGGAGCAGATGATTTGTGACTACACAAGACAGTGTGTTGGGGCTACTGTATTAC 1389
Qy 1537 aaactgcagggcctctcccttggacagcctgggctgtttctccaaagtgtcgaagcct 1596
Db 1390 AAACCTCTTGTGAATTCAAGCTGATGTGTGGGTGCTTCCACAAATGTGCGAAACCT 1449
Qy 1597 tgtagtgtatcaactacaactctctgcggctactcacggtggccatctgtgaagtc 1656
Db 1450 TGCAAAATGACAGAAATACCAGCTCTCAGCTGGATACTCTCGCTGGCTTCTGTGCTCA 1509
Qy 1657 caggattgatatctcgagatgtgtctctgcagaacaataacactattacaacaacaaaga 1716
Db 1510 GAGGATTTGGGTTTTTACATGCTTTACACACAGAACAAATAACAATAATCAATCAAGAGG 1569
Qy 1717 aacggagttgcaagctcaacatcttctcaaggagctgaactataaaactaatctcgag 1776
Db 1570 ATGGAGTTTGCAAAAGTGAATATCTTTTGGAGGAATGGAACCTACAAAGCAATGGGAG 1629
Qy 1777 tctctctctgtcacagatggttcagcctcctgtccaaacctgggcagccagtggaacctgtgg 1836
Db 1630 TCACCAGCCTTTACAGTAGTACCCCTCTTGTCTCAGCTTTGGGAATCAGTGGAGTCTCTGG 1689
Qy 1837 ttgtgctgtccgtctctctgtgtggagatgaggagcgtcatcttcgaacctcttggtc 1896
Db 1690 TTTGGATCTCTGCTCTATCAGTGTGAGGAGCTTGCAGAGCTGATTTCTGGATTTCACTGTC 1749
Qy 1897 atcacacttctcatgctgactacgcgggttccgcggagccggtactgg 1941
Db 1750 ATCACTTTTCATCTGCCCTTCCGCTGGTTCCGCTCCAGCGGTGG 1794

RESULT 13

AF232715 1003 bp mRNA NAM 01-MAR-2000
LOCUS Ovis aries amiloride-sensitive sodium channel alpha subunit mRNA,
DEFINITION partial cds.
ACCESSION AF232715
VERSION AF232715.1 GI:7110531
KEYWORDS sheep.
SOURCE Ovis aries
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
Caprinae; Ovis.
REFERENCE 1 (bases 1 to 1003)
AUTHORS Barley,J., Syrris,P. and Walters,D.
TITLE Expression of sheep lung amiloride-sensitive epithelial sodium
channel alpha subunit gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1003)
AUTHORS Barley,J., Syrris,P. and Walters,D.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2000) Child Health, St. Georges Hospital Medical
School, Cranmer Terrace, Tooting, London SW17 0RE, UK
FEATURES
source
1..1003
/organism="Ovis aries"
/db_xref="taxon:9940"
/tissue_type="lung epithelium"
CDS
<1..>1003
/note="OENac"
/codon_start=1
/product="amiloride-sensitive sodium channel alpha
subunit"
/protein_id="AAF36982.1"
/db_xref="GI:7110532"
/translation="KDWKIGFQLCNQNKSCDFYQTYSSGVDAVREWRPHYINILSR
RODTSPLEEDVLGKFIETFRNQDSCNEANYSHFHPMYGNCYTTNDKNSNLWISS
VFGVNNGLSLTRTEQNDIFPLLSVTGARVMYHREDPEAFMDGAFNLRPGVKVTSIS
MSKEAVDRLLGGYDGDCTKNGSEVPVGNLYNTTVQOVCIHSCFSEIKKECGCAVIFY
PRPEGVEFCDYKKHNSWGVCYYKLQDAFSSDRLLGCTFKCRKPCSVTIYKLASYSOMP
SVTSODWVFQMLSRNNYTIKNKRDGVAKLNTFFKELNYKSNSSEFSVMTVLLSLNG
"

BASE COUNT 244 a 285 c 255 g 219 t
ORIGIN
Query Match 24.1%; Score 751.8; DB 3; Length 1003;
Best Local Similarity 85.0%; Pred. No. 1.7e-170;
Matches 853; Conservative 0; Mismatches 147; Indels 3; Gaps 1;
Qy 820 aaggactggaagatcgcttccaactgtgcaaccagaaacacacagactgtttctaccag 879
Db 1 AAGGACTGGAAGATCGCTTCCAGCTGTGCAACCAACAAATCGGACTGCTTCTACCAG 60
Qy 880 acatactcctctgggtggatgcagtgagggagtggtaccgcttccattacaacatt 939
Db 61 ACATATCTGTCAGGGTGGATGCTGTGAGGAGTGTGTATCGCTTCCACTATATCAACATT 120
Qy 940 cgtgtccagactg---tcggacacctgcgcgtctctagaggaagaagccctgggcaacttc 996
Db 121 CTGTCCAGAGGAGACAGGACACCTCTCCCTCGCTGGAGGAGGAGCTGTGGGCAAAATTC 180
Qy 997 attctcaactgtgtcgttcaaccaggccccctgcaaccaggcgaattattccaagtctccac 1056
Db 181 ATCTTACCTGCCGTTTCAACAGGACTCTCTGCAATGAGGCGAATTACTCTCACTCTCCAC 240
Qy 1057 caccctcatgaggggaactgtacactttcaatgacaagaacaactccaatctctgagatg 1116
Db 241 CACCCGATGTACGGAACACTGCTACACTTTTCAACGACAAAGAACAGCTCCACCTCTGGATA 300
Qy 1117 tcttccatgctgagtcacaatggtttgtccctgacactgcgcagcagacagaatgac 1176
Db 301 TCCTCCGTGCCGGGTCAACACGGCCTGTCCCTGACACTGCGCAGACAGACAAATGAC 360

QY	1177	ttatccccctgctctccacagtcaggggcccagggatggtgatggtcagtgatgag	1236
Db	361	TTATCCCACTGCTGTCCACGGTGACGGGGCGGGTGATGTCATGAACGGGATGAG	420
QY	1237	ctgctcttatgatgatggtggtctcaactgaactgagcgtggtgagacccctcaagt	1296
Db	421	CTGTCCTTTATGATGATGCTGCTTTAATCTGCGGCCGGCGTGAGACCTCCATCAGC	480
QY	1297	atgagaagaacccctggacacccctcgaggaataatgaagcactgtactgagaatggt	1356
Db	481	ATGAGTAAGGAACCGTGGACACTTGGGGGTGACTATGGCGACTGCACCAAGAATGGC	540
QY	1357	agcagatccccggtcaagaacctttacccttccaagtatacacagcaggtgtgctattcac	1416
Db	541	AGCGAGGTCCCGTGGGGAACCTTTATATACCACGTACACGACGAGGTGTGCATCCAT	600
QY	1417	tctgtctccagagagaacatcatcaagaagtgtggtgtgtcctacatctctccacctaa	1476
Db	601	TCGTGCTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGTGCCTTACATCTTCTACCCACGG	660
QY	1477	ccaaaggaggtgagttctgtgactaccgaaagacagagctctctgggctatttctattat	1536
Db	661	CTGAGGGGTAGATGCTGTGACTACAGAAGCATAAATTTTGGGGCTACTGCTACTAT	720
QY	1537	aaactgcagggcctctctctgtggacagcctgggtgtgttctccaaagtgtcggaagcct	1596
Db	721	AAGCTCAGGATCCCTTTCCTCAGACCGCGCTGGCTGTTCACCAAGTGC CGGAGGCCA	780
QY	1597	tgtagtgtgatcaatacaaaactctgtgcggctactcaogtggccatctgtgaagtcc	1656
Db	781	TGCAGTGTGACCATCTACAAAGCTCTCTGCCAGTTACTCAGATGGGCCCTCCGTGACATCC	840
QY	1657	caggattgatcttcagatgctgtcctctgcagaaacattacactatttaacaacaaaaga	1716
Db	841	CAGGACTGGCTCTCCAGATGCTGTCCCGGCAGAACAACTACACCATCAAGAAACAAAGG	900
QY	1717	aacggagttgcaagctccaacatctcttcaaggagctgaactataaaactaatctcgag	1776
Db	901	GATGGAGTTCACCAACTCAACATCTTCTCAAGGAGCTGAACATCAAAATCTAATCTCTGAG	960
QY	1777	tctctctgtcacagtgtgcagcctctcttccaaacctgggca	1819
Db	961	TCCTCCTCAGTCACGATGGTCAACCTCTCTGTCCAACCTGGGCA	1003
RESULT 14			
GGU62902			
LOCUS	2321 bp	mRNA	VRT
DEFINITION	Gallus gallus amiloride-sensitive epithelial sodium channel alpha subunit (alpha cEnAc) mRNA, complete cds.		
ACCESSION	U62902		
VERSION	U62902.1	GI:1497998	
KEYWORDS	Chicken.		
SOURCE	Gallus gallus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
REFERENCE	1	(bases 1 to 2321)	
AUTHORS	Killick,R. and Richardson,G.		
TITLE	Isolation of chicken alpha EnAc splice variants from a cochlear cDNA library		
JOURNAL	Biochim. Biophys. Acta 1350 (1), 33-37 (1997)		
MEDLINE	97157073		
REFERENCE	2	(bases 1 to 2321)	
AUTHORS	Killick,R. and Richardson,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-JUL-1996) School of Biology, Sussex University, Falmer, Sussex BN1 9QJ, UK		
FEATURES	Location/Qualifiers		
source	1..2321		
	/organism="Gallus gallus"		
	/strain="Isa Brown"		
	/db_xref="taxon:9031"		

gene	/tissue_type="cochlea"			
	89..2002			
	/clone="2.5kb transcript"			
CDS	89..2002			
	/gene="alpha cENaC"			
	89..2002			
	/gene="alpha cENaC"			
	/note="contains two hydrophobic domains, a cysteine rich domain, and a proline rich region"			
	/codon_start=1			
	/product="amiloride-sensitive epithelial sodium channel alpha subunit"			
	/protein_id="AAB50550.1"			
	/db_xref="GI:149799"			
	/translation="MGTASRGVSKAEKMEPEKTRQCKOETEQOQEKDEREGLIEFYGSYQDVQFQCSNNTTIHGAIRLVCSKKMKMTAFWSVLFILFTEGLMWQFGLYRYFSPVNLNLNSDRILFFPAVLTLNPIRYSAIRKKLDELQDHTQLDLQDLYNMSLARSDSAQFSHRSTRSLHVRHPLRQKDNLVSLPENSVDKNQWKIGFVLCSENNEDFHQYSSGVDVAREWYSFYINILAQMPDARDLDEDFENFIYACRNEATCDKANTFHHPILYGNCTYFNDSNSSLTSLPGINNGLSLVVTRTEQNDIPLILSTVTGARVMYHDONEPAFMDGDFNVRPGIETISIMRKEMTERLGSSYSDCTEDGSDVPQNLYSRYTEQVCIKSCFQLNMKRSCAYEYFPLPDGAEYCDYTKHWANGCYCYKLLAEFKADVIGCFHKRRPKCKMETQLSAGYSRMPSSAVSDEWYFMLSQQKNYINITSKRNGVAKNIFEEWNYKTINGESPATVVVTLISQLGNOWSLWFGSSVLSVMELAEILIDFTVITFLAIFRFRSKQWHSSPAPPNSHDNTAFQDEASGLDAPHRTVEAVVTTLPSYNSLPCGPSKDGEGTGLE"			
BASE COUNT	584 a	584 c	578 g	575 t
ORIGIN				
Query Match	24.0%; Score 747.4; DB 4; Length 2321;			
Best Local Similarity	66.4%; Pred. No. 2.1e-169;			
Matches 1145; Conservative	0; Mismatches 556; Indels 24; Gaps 4;			
QY	217	gggaagggggaacaaogtgaagagcagggctggccggcggaacccctcagaccccgagcag	276	
Db	125	GAGAAGATGCCAGAAGGGGAGAAGACGAGGACGTGCAAGCAGGAAACGGAGCAGCAG	184	
QY	277	cccacagagagagagagcactgattgaattccacgcctcctcaccggagctcttcag	336	
Db	185	AAGAAGATGAGAGGGAGGCCCTCATCGAATTCATGGCTCTCTACCAGGACGTTCCAG	244	
QY	337	ttcttctgcaacaacaccaccatccacggggccatccgcctcgtgtgtctcacaacacacac	396	
Db	245	TTCTTCTGTAGCAACACGACCATCCACGGAGTATTTCGCCTGGTGTCTCTAAAAAGAAC	304	
QY	397	cgcataagacgcctctctggcggtgctgtgctgtgcaccttcggcatgactgactgg	456	
Db	305	AAGATGAAGACGGCCTTCTGTGCTGTCTCTTTCATTCACCTTCGGCTTAATGATTGG	364	
QY	457	cagttgccttgcctgttcgagagagtagctcagctaccagtgagcctcaacatcaacctc	516	
Db	365	CAGTTGGAAATCCTCTACAGGGAGTACTTCAGCTACCTGTCAACCTCAACCTCAACCTA	424	
QY	517	aattcagacaagctggtcttcctcctgcctgcctgctgtgcaccttaacatcctacagatac	576	
Db	425	AACTCGGATAGGTGACTTTCCCTGCGAGTGACGCTGTGCACACTCAACCCATACAGGTAC	484	
QY	577	actgaaataaagagagctggaaagagctgagacgcatcacggagacagagcgtttttgac	636	
Db	485	AGTGCCATTTCGGAAGAGCTGGAGAACTGGACCAATACACATCAGACACTGTTGGAC	544	
QY	637	ttgtacaaatacaactcttctacactcgcagcgtggcgccgagccgcgagctccgcg	696	
Db	545	CTTTATGACTATACATGTCCCTGGCACGAAGTGACGGGTGACGGCTGCTACATAGG	604	
QY	697	gacctcctgggtgttttcccgacccccctgcagcgctgcgcactccacctccgcccctac	756	
Db	605	CGCACTCGAGGAGCCCTGCTCCATCATGTTTCAGCGCCATC-----CTCTGC	650	
QY	757	tcggcgcgacggcgcgagcgggtcttccaggtacgcgcagacaaacatccccaaagtggac	816	
Db	651	GGAGGCGAAGACGGGATAACTTGGT----CAGCTTGCCAGAAACACGTCCCTCGGTGGAC	706	

[illegible]

Db 245 TTCCTTCTGTACACACAGCACCATCCACGGAGCTATTTCGCCCTGGTGCTCTAAAGAAC 304
Qy 397 cgcataagacgagcctcttgccgagtgctgtggctgtgcaccttcgcgcagatgactgg 456
Db 305 AAGATGAAGACGGCCTTCTGCTCTGCTCTCATTTCTCACCTTCGGCTTAATGTATTGG 364
Qy 457 cagtgcccttgctgttgagagagtaacctagctaccocagtgagcctccaaacatacctc 516
Db 365 CAGTTTGGAAATCCCTACAGGAGTACTTTCAGCTACCTACCTGTCAACCTCAACCTA 424
Qy 517 aattcacagaagctggctctccctccgctcaactgtctgcacccttaactcttacagatac 576
Db 425 AACTCGGATAGGCTGACTTCCCTGCGAGTGACGTGCGACATGCGACATCAACCCATACAGGTAC 484
Qy 577 actgaataaagaggagctggaagagctgacccgcacacagcgacagagcagccttttgac 636
Db 485 AGTCCATTCGGAGAGAGCTGAGAGCAATGGACCAATCACACATCAGACACTGTGGAC 544
Qy 637 ttgtacaataaactcttctacactcgcagagctggggccgcagccgcagctccgcg 696
Db 545 CTTTATGACTATAACATGTCCCTGGCAGAAAGTGACGGGTGACGGCAGTTCTCACATAGG 604
Qy 697 gacctctgggtgttctcccgacccccctgcagcgctgcgaactccacactccgcctac 756
Db 605 CGCACCTCGAGGAGCCCTGCTCCATCATGTTTCAGCGCCATC-----CTCTGC 650
Qy 757 tccggccgcagcgcgcgagcgggtcttccagcgtacgcagacaacaatccccaagtggac 816
Db 651 GGAGCGAGAGCGGGATACTTGGT----CAGCTTGGCAGAAACAGTCCCTCGGTGGAC 706
Qy 817 cgggaagactggaagatcggtcttccaaactgtgcaaccagacaacaatcagagctttctac 876
Db 707 AAGAAGGACTGGAATAATGGCTTGTCTGTGTCAGTGAACAAATGAGGATGTTTCCAT 766
Qy 877 cagacatactcctctg999tgatgcagtgaggagtggtaccgcttccattacatacaac 936
Db 767 CAGACGTACTCTCAGGGGTGGATGCGATGCGGGAATGGTACAGCTTTTCACTACATCAAT 826
Qy 937 attctgtccagactgtcggacacctcgcgcctctagagaagaagccctgggcaacttc 996
Db 827 ATCTGGCAGAGATGCTGATGCC---AAGACCTGGATGAGTGTGACTTTGAGAAATTT 883
Qy 997 attctacctgtcgtcttaacacaggcccttgcacaggcggaattatttccaaagtccac 1056
Db 884 ATTTATGCTTGCCGCTTCAATGAAGAACATGTGACAAGGGGAATTACACTCACTTCCAC 943
Qy 1057 caccctatgtacgggaactgtacactttcaatgacaagaacaactccactctctgagatg 1116
Db 944 CATCCCTTGTATGGGAATGCTGTACACCTTCAATGACAATAGCAGC---AGCCTGTGGACA 1000
Qy 1117 tctccatgcctgtgagtgacaactgtttgtccctgacactgcgcacagacagaaatgac 1176
Db 1001 TCCCTCTGCTGGGATCAATAATGGGCTCTCTGTGGTGGCCAGCCGACGAGAATGAT 1060
Qy 1177 ttcattccccctgtctccacagtgacggggccaggggtgagtgtgcattggtcaggatgag 1236
Db 1061 TTCAATCCCTCTGTGTTCAACGGTAACTTGGACCCAGGGTCAAGTCCAGATCAGAATGAG 1120
Qy 1237 cctgccttattgagatgagtggtgttcaacttggagcctggcgtggagacctccatcagt 1296
Db 1121 CCAGCCTTCATGGATGATGGGGTTTCAATGTGCGTCCAGGTATGAGACCTCCATCAGC 1180
Qy 1297 atgagaaaggaagccctggacagcctcggaggaataatcaggcgactgtactgagaatggt 1356
Db 1181 ATGAGGAAGGAGATGACTGACGCTTGTGGGGCAGTTACAGCGACTGTACAGAGGATGGC 1240
Qy 1357 agcagatccccggtcaagaaccttacccttccaaagtatacacagcaggtgtgcatcac 1416
Db 1241 AGCGATGTGCCAGTGCRAAATCTGTACTCATCTCGCTACACTGAACAGGCTGCATTCGT 1300
Qy 1417 tccctgtccaggagaacatgatcaagaagtgtggtgtgcctacatcttctaccctaaag 1476
Db 1301 TCCTGCTTTACGCTCAACATGGTAAGCGCTGCTCCTGTGCGCTATTACTTTTATCCCTTA 1360

Qy 1477 cccaagggagttgagttctctgtgaactaccgaaagcagagctcctcgggctattgctattat 1536
Db 1361 CCTGATGGAGCAGAGTATTGTGACTACACAAGACAGCTAGCTTGGGGCTACTGCTATTAC 1420
Qy 1537 aaactgcaggcgcccttctccttgagacagcctgggctgtttctcocaagtgtcggaagcct 1596
Db 1421 AAACCTCCTTGTCTGAATTCAAAAGCTGATGTCTGGGCTGCTTCCACAAATGTGGAAAGCT 1480
Qy 1597 tgtagtgatcaactacaactctctcggcgtactcagcgggtgcceatctgtgaagtc 1656
Db 1481 TGCAAAATGACAGAATACCAGCTCTCAGCTGGATACTCTCGCTGGCCTTCTGCTGTCTCA 1540
Qy 1657 caggattgatatcttcgagatgctgctccttcagacaataatcacactattacaacaaaaga 1716
Db 1541 GAGGATTGGGTTTTTTTACATGCTTTTCAACAACAGACAATAATACATATACATCCAGAGG 1600
Qy 1717 aacggagttgcaagctcaacatcttctcgaagagctggaactataaactaatctcgag 1776
Db 1601 AATGAGTTGCAAAAGTGAATATCTTTTTTGGAGGAATGGAACACAGCAATGGGGAG 1660
Qy 1777 tctcctctgtcaagatggttcagcctcctgtccaaacctggcgagccagtgagcctgtgg 1836
Db 1661 TCACCAGCCTTTACAGTAGTACCCCTCTTGTCTCAGCTTGGGAATCAGTGGAGTCTCTGG 1720
Qy 1837 ttggctcgtcgtgctctctgtgtgagatggcgagcgtcatcttcgacctcctgctc 1896
Db 1721 TTTGGATCCTCTGTCCCTATCAGTGTATGGAGCTGTCTGAGCTGATTTCTGGATTTCACTGTC 1780
Qy 1897 atcacacttctcatgctgctacgcgggttcgggagccgggtactgg 1941
Db 1781 ATCACTTTTCATCCTGGCCCTTCCGCTGGTTCCGCTCCAAGCAGTGG 1825

Search completed: September 10, 2000, 23:02:51
Job time: 5759 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 22:15:07 ; Search time 97.02 Seconds
(without alignments)
8038.020 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcaccacgcgtcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0
Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	45	1.4	45	1 T99077	Probe Ena-A7 for a
C 2	45	1.4	45	1 T99078	Probe Ena-A9 for a
C 3	38	1.2	75	1 T99072	Alpha epithelial s
C 4	38	1.2	960	1 Q03293	Recombinant DNA en
C 5	38	1.2	1276	1 V29126	HPTP-1 coding sequ
C 6	38	1.2	3557	1 N60801	Human pro-growth h
C 7	38	1.2	3621	1 N60847	Human pre-prolacti
C 8	37	1.2	495	1 T74484	Pilosulin 2 precu
C 9	37	1.2	606	1 V73439	Flea saliva protei
C 10	37	1.2	732	1 V59566	Flea saliva protei
C 11	37	1.2	732	1 V59566	Human secreted pro
C 12	37	1.2	970	1 N50414	Sequence encoding
C 13	37	1.2	1001	1 T73851	Full length 3' UTR
C 14	37	1.2	1889	1 V69433	Human edg-6 cDNA #
C 15	37	1.2	2158	1 V25156	Alpha-GalNac from
C 16	37	1.2	2702	1 V55872	Plant acetolactate
C 17	36	1.2	340	1 V51737	DNA encoding a hum
C 18	36	1.2	378	1 V90045	EST clone CW1550.
C 19	36	1.2	433	1 Q23343	Murine Cp-10. New
C 20	36	1.2	437	1 X22245	Human secreted pro
C 21	36	1.2	505	1 X04408	Human secreted pro
C 22	36	1.2	558	1 T64563	Mouse thymus and a
C 23	36	1.2	581	1 T63172	Trypanosoma cruzi
C 24	36	1.2	582	1 T64562	Human thymus and a
C 25	36	1.2	602	1 X20419	Human secreted pro
C 26	36	1.2	684	1 Q72956	Clone pSP65-xa1-11
C 27	36	1.2	687	1 X30351	DNA encoding a hum
C 28	36	1.2	742	1 Q22005	Sequence encoding
C 29	36	1.2	780	1 T59912	Human transcriptio
C 30	36	1.2	794	1 N92416	Sequence of the S3
C 31	36	1.2	808	1 X04314	Human secreted pro
C 32	36	1.2	886	1 X51754	DNA encoding a hum
C 33	36	1.2	893	1 Q92529	P. communis (pear)

34	36	1.2	894	1 V99723	Human adult retina
35	36	1.2	921	1 X27330	Human secreted pro
36	36	1.2	921	1 X51728	DNA encoding a hum
37	36	1.2	971	1 Q73760	RP-8 Programmed ce
38	36	1.2	971	1 T35704	Coding sequence fo
39	36	1.2	990	1 X51748	DNA encoding a hum
40	36	1.2	1008	1 X37470	Human secreted pro
41	36	1.2	1029	1 V45444	Human chemokine 2S
42	36	1.2	1037	1 X30150	Human secreted pro
43	36	1.2	1046	1 Q10377	Plasmid pMG3C9 use
44	36	1.2	1060	1 X37464	Human secreted pro
45	36	1.2	1089	1 X37460	Human secreted pro
46	36	1.2	1095	1 X30366	DNA encoding a hum
47	36	1.2	1113	1 T75765	P. tetragonolobus
48	36	1.2	1117	1 V59118	Nucleotide sequenc
49	36	1.2	1142	1 V43611	Human secreted pro
50	36	1.2	1152	1 V62754	Human secreted pro
51	36	1.2	1259	1 Q23028	Clone W264 encodin
52	36	1.2	1270	1 Q66589	Wheat leaf ADP-glu
53	36	1.2	1273	1 X06777	Human sperm-specif
54	36	1.2	1284	1 Q15361	NP-YB. Expression
55	36	1.2	1325	1 V80740	Human secreted pro
56	36	1.2	1332	1 V40647	Human transaldolas
57	36	1.2	1332	1 X02709	Human transaldolas
58	36	1.2	1338	1 Q65607	Rabbit zona pelluc
59	36	1.2	1338	1 V64789	Rabbit zpc cDNA. I
60	36	1.2	1389	1 X27067	Original S. tubero
61	36	1.2	1395	1 Q75359	Gibberellin-20-ox
62	36	1.2	1425	1 X19486	Human secreted pro
63	36	1.2	1441	1 T90173	Oil seed rape cyst
64	36	1.2	1442	1 X06779	Human sperm-specif
65	36	1.2	1521	1 V73000	Human foetal brain
66	36	1.2	1588	1 Q23029	Clone W282 encodin
67	36	1.2	1602	1 T43204	Human mitogen-acti
68	36	1.2	1602	1 X07066	Human mitogen acti
69	36	1.2	1623	1 T14925	T cell replacing f
70	36	1.2	1623	1 V64062	Plasmid pSP6K-mTRF
71	36	1.2	1632	1 Q71341	Human basigin I im
72	36	1.2	1638	1 T06480	Cystathionine gamm
73	36	1.2	1704	1 V34509	Human TNF receptor
74	36	1.2	1704	1 V83763	Human tumour necro
75	36	1.2	1705	1 V08856	Gene encoding huma
76	36	1.2	1705	1 V84616	Human secreted pro
77	36	1.2	1739	1 Q27539	Tyrosine Kinase re
78	36	1.2	1808	1 T30394	Oligodendrocyte-my
79	36	1.2	1952	1 Q39085	Xr1 coding sequenc
80	36	1.2	1977	1 X27383	Human secreted pro
81	36	1.2	2082	1 V41257	Mouse neuronal PAS
82	36	1.2	2119	1 Q43853	Sequence of part o
83	36	1.2	2140	1 V42316	Macadamia integrif
84	36	1.2	2182	1 V00731	Pea plastidial pho
85	36	1.2	2353	1 V45981	A. thaliana sterol
86	36	1.2	2381	1 Q68910	Human c-myc far up
87	36	1.2	2384	1 Q68909	Human c-myc far up
88	36	1.2	2420	1 Q46672	p.mysa cDNA clone.
89	36	1.2	2496	1 V63189	cDNA from clone cr
90	36	1.2	2567	1 V44889	Coding sequence fo
91	36	1.2	2779	1 V59691	Human secreted pro
92	36	1.2	3076	1 V73003	Human adult brain
93	36	1.2	3238	1 Q83532	Sequence encoding
94	36	1.2	3395	1 V40547	Homo sapiens secre
95	36	1.2	3451	1 V02308	Cell membrane prot
96	36	1.2	3684	1 X21560	Human N-arginine d
97	36	1.2	3784	1 Q05326	Sequence encoding
98	36	1.2	3820	1 Q03517	Plasmid pfgH1 enco
99	36	1.2	3848	1 X27277	Prostate-tumour de
100	36	1.2	3850	1 N81634	Sequence of new pl
101	36	1.2	3851	1 T79857	Yellowtail cuna DN
102	36	1.2	3851	1 T90019	Yellowtail/fin tu
103	36	1.2	3989	1 V84428	Human secreted pro
104	36	1.2	4253	1 V40888	Coding sequence of
105	36	1.2	4586	1 T96838	Intron 8 of human
106	36	1.2	4874	1 V49807	Human ADNF-III cDN

107	36	1..2	4978	1	Q74281	PLA2-I receptor en	180	35	1..1	1036	1	V84486	Human secreted pro
108	36	1..2	4978	1	T02986	Bovine phospholipase	181	35	1..1	1040	1	V02818	Human RHAMM genom
109	36	1..2	5668	1	X27262	Prostate-tumour de	182	35	1..1	1066	1	Q49943	Human anti-HBs lig
110	36	1..2	6254	1	V40887	Coding sequence of	183	35	1..1	1069	1	V84474	Human secreted pro
111	36	1..2	237326	1	V57903	Hereditary haemoch	184	35	1..1	1071	1	X20414	Human secreted pro
112	35	1..1	57	1	V27940	Human GAI3 primer	185	35	1..1	1097	1	V44860	Clone BG142_1 codi
113	35	1..1	70	1	T17035	Human mitochondria	186	35	1..1	1128	1	T86027	Meripilus giganteu
114	35	1..1	75	1	T17036	Human mitochondria	187	35	1..1	1180	1	T86347	Human Apo-2DCR cDN
115	35	1..1	85	1	V32418	Homo sapiens clone	188	35	1..1	1181	1	V59803	Human secreted pro
116	35	1..1	90	1	V37789	Homo sapiens clone	189	35	1..1	1196	1	V20812	Human haematopoiet
117	35	1..1	106	1	V00430	Analytical solid p	190	35	1..1	1212	1	V59886	Human secreted pro
118	35	1..1	106	1	V11611	3' fragment of clo	191	35	1..1	1215	1	X00620	Human secreted pro
119	35	1..1	160	1	T61985	Homo sapiens adult	192	35	1..1	1225	1	T33325	Human secreted pro
120	35	1..1	179	1	V11620	Detection probe fo	193	35	1..1	1248	1	V58644	American elm chiti
121	35	1..1	204	1	T61988	Homo sapiens adult	194	35	1..1	1248	1	V61249	Prostate tumour sp
122	35	1..1	207	1	Q50399	Detection probe fo	195	35	1..1	1261	1	X04382	cDNA sequence of p
123	35	1..1	259	1	V00423	GPf - 1E, part of	196	35	1..1	1276	1	T42071	Human secreted pro
124	35	1..1	263	1	T85891	3' fragment of clo	197	35	1..1	1285	1	Q86744	Human nuclear prol
125	35	1..1	283	1	V89992	Malassezia fungus	198	35	1..1	1302	1	Q11112	TRMP-3 metalloprot
126	35	1..1	308	1	V02985	EST clone CW762, N	199	35	1..1	1305	1	V26270	MHC class II antig
127	35	1..1	308	1	V11421	Human BF245_1 3' p	200	35	1..1	1305	1	V26270	Trichomonas vagina
128	35	1..1	317	1	Q55664	Human secreted pro	201	35	1..1	1307	1	V63176	Trichomonas vagina
129	35	1..1	470	1	V58579	AG1 toxin, New pep	202	35	1..1	1322	1	X04402	cDNA from clone dx
130	35	1..1	470	1	V61312	Prostate tumour sp	203	35	1..1	1328	1	V58754	Human secreted pro
131	35	1..1	509	1	V58672	cDNA sequence of p	204	35	1..1	1337	1	X04345	Human secreted pro
132	35	1..1	509	1	V61342	Prostate tumour sp	205	35	1..1	1355	1	N80457	Human secreted pro
133	35	1..1	553	1	V31993	Extended cDNA sequ	206	35	1..1	1361	1	X51714	Sequence of cDNA e
134	35	1..1	571	1	V04715	BS106 polynucleoti	207	35	1..1	1380	1	T51051	DNA encoding a hum
135	35	1..1	572	1	V44301	Human secreted pro	208	35	1..1	1386	1	V11855	Human anine recept
136	35	1..1	579	1	Q14800	Human secreted pro	209	35	1..1	1391	1	T07329	Homo sapiens Tub I
137	35	1..1	582	1	X37502	Human chorionic go	210	35	1..1	1423	1	N90249	Human calcium sign
138	35	1..1	588	1	V84464	Human secreted pro	211	35	1..1	1426	1	T33115	Cucumber mosaic vi
139	35	1..1	588	1	X07427	Human secreted pro	212	35	1..1	1426	1	T72272	Human placental ri
140	35	1..1	600	1	X30350	Homo sapiens secre	213	35	1..1	1448	1	X37462	Cucumber mosaic vi
141	35	1..1	603	1	T91169	DNA encoding a hum	214	35	1..1	1467	1	V82457	Human secreted pro
142	35	1..1	630	1	T45865	Secreted cytokine	215	35	1..1	1493	1	Q28411	Triticum sp. cyste
143	35	1..1	631	1	V82801	cDNA encoding C. h	216	35	1..1	1503	1	V29523	Glia activating fa
144	35	1..1	635	1	Q75445	ATG-1120 (allograf	217	35	1..1	1512	1	T72172	Homo sapiens PSP1
145	35	1..1	635	1	T38390	Murine glycosylati	218	35	1..1	1522	1	V46309	Alzheimer's diseas
146	35	1..1	643	1	V84539	Murine glycosylati	219	35	1..1	1522	1	V22704	Human secreted pro
147	35	1..1	669	1	N90103	Human secreted pro	220	35	1..1	1525	1	X00679	Mouse recombinase
148	35	1..1	685	1	X27366	Human prealbumin c	221	35	1..1	1527	1	Q35200	Human secreted pro
149	35	1..1	685	1	X37398	Human secreted pro	222	35	1..1	1534	1	T15995	DHP-1 polypeptide
150	35	1..1	687	1	X30410	Human secreted pro	223	35	1..1	1538	1	V84593	Human interleukin-
151	35	1..1	688	1	T16856	DNA encoding a hum	224	35	1..1	1540	1	V82787	Human secreted pro
152	35	1..1	712	1	V59811	Integrin subunit b	225	35	1..1	1547	1	X33448	Clone gq38_1 isola
153	35	1..1	713	1	Q48773	Human secreted pro	226	35	1..1	1554	1	X16675	Oryza sativa L. pi
154	35	1..1	713	1	Q49210	Aphrodite, DNA a	227	35	1..1	1573	1	V43607	Xenopus WA545 prot
155	35	1..1	745	1	V82783	Hamster Aphrodite	228	35	1..1	1576	1	Q49944	Human secreted pro
156	35	1..1	755	1	Q03970	Clone dn809_5 isol	229	35	1..1	1607	1	T85878	Human anti-HBs hea
157	35	1..1	767	1	T15821	Clone P25b New pro	230	35	1..1	1613	1	X26196	Malassezia fungus
158	35	1..1	767	1	T44470	TBE virus strain N	231	35	1..1	1621	1	V58584	DNA sequence of gl
159	35	1..1	801	1	N91234	Tick-borne encephal	232	35	1..1	1621	1	V61199	Prostate tumour sp
160	35	1..1	831	1	V84475	cDNA clone 8F3 of	233	35	1..1	1637	1	V84578	Full length cDNA s
161	35	1..1	832	1	X30363	Human secreted pro	234	35	1..1	1642	1	V16096	Human secreted pro
162	35	1..1	857	1	Q46850	DNA encoding a hum	235	35	1..1	1642	1	V40507	Tyrosine kinase as
163	35	1..1	857	1	T14564	Recombinant human	236	35	1..1	1651	1	Q80218	Human secreted pro
164	35	1..1	863	1	T91302	Human milk kappa-c	237	35	1..1	1670	1	V46310	Human secreted pro
165	35	1..1	877	1	V014635	Human HI075-1 secr	238	35	1..1	1686	1	T91311	Human J143-1 secre
166	35	1..1	885	1	N90091	Human cytoplasmic	239	35	1..1	1693	1	V69206	Canine ICAM-R poly
167	35	1..1	907	1	V59695	Gene 519 cDNA New	240	35	1..1	1720	1	X27439	Human secreted pro
168	35	1..1	940	1	T58281	Human secreted pro	241	35	1..1	1733	1	V81394	Human tumour antig
169	35	1..1	941	1	V84586	Arabidopsis STZ po	242	35	1..1	1755	1	Q71873	Sequence coding hu
170	35	1..1	954	1	X37477	Human secreted pro	243	35	1..1	1767	1	X51749	DNA encoding a hum
171	35	1..1	971	1	X37469	Human secreted pro	244	35	1..1	1781	1	Q46991	cDNA clone encodin
172	35	1..1	978	1	X51732	Human secreted pro	245	35	1..1	1781	1	V19328	Human ICAM-4 cDNA
173	35	1..1	980	1	X37466	DNA encoding a hum	246	35	1..1	1781	1	V11657	Human ICAM-R cDNA
174	35	1..1	981	1	V60576	Human secreted pro	247	35	1..1	1781	1	V56348	Human ICAM-R cDNA
175	35	1..1	986	1	X51735	ERAB protein codin	248	35	1..1	1781	1	V54926	cDNA encoding huma
176	35	1..1	996	1	Q89606	DNA encoding a hum	249	35	1..1	1781	1	V69125	Interleukin adhe
177	35	1..1	1013	1	Q12431	Human membrane pre	250	35	1..1	1781	1	X21855	Human ICAM-R codin
178	35	1..1	1023	1	T70132	Fragment D of urat	251	35	1..1	1813	1	V70895	cDNA encoding an a
179	35	1..1	1024	1	V84444	Max-interacting pr	252	35	1..1	1921	1	T08610	Yeast heterogene e

253	1.1	1965	1	Q11815	HT7 membrane prote	326	1.1	5769	1	V15004	Receptor protein t
254	1.1	1971	1	X10156	Sequence encoding	327	1.1	5894	1	T85474	hABC3 cDNA sequenc
255	1.1	1984	1	Q52252	Protein PRO271 cDN	328	1.1	5894	1	V16307	hABC3 encoding huma
256	1.1	2001	1	T59700	PTH-like peptide D	329	1.1	5975	1	N91778	Rabbit skeletal mus
257	1.1	2025	1	V59809	Human secreted pro	330	1.1	5975	1	Q87978	Rabbit skeletal ca
258	1.1	2045	1	V34297	Human secreted pro	331	1.1	5975	1	T70228	Rabbit calcium cha
259	1.1	2065	1	T58164	Rice ADP-glucose p	332	1.1	5975	1	T96811	Rabbit skeletal ca
260	1.1	2085	1	T29062	BRCA1 modulator pr	333	1.1	6476	1	X24302	Murine mWRN gene (
261	1.1	2094	1	X19489	Human secreted pro	334	1.1	6525	1	V16345	cDNA encoding full
262	1.1	2116	1	Q35987	Tomato hsp80 cDNA	335	1.1	6644	1	X33181	Base sequence of t
263	1.1	2132	1	X33816	Coding sequence fo	336	1.1	7372	1	X33182	Base sequence of t
264	1.1	2144	1	T23538	Homo sapiens FSP1-	337	1.1	7797	1	X33180	Cowpox virus bsr f
265	1.1	2152	1	T22002	Human MED1 endonuc	338	1.1	7859	1	Q12154	Shuttle vector pSW
266	1.1	2158	1	T15228	Tumor necrosis fac	339	1.1	7859	1	Q44265	PSW6 for expressio
267	1.1	2158	1	T94631	TNF-R1-DD ligand p	340	1.1	7984	1	Q25185	PSW6 expression ve
268	1.1	2171	1	V73006	Human foetal brain	341	1.1	7996	1	X33184	Base sequence of t
269	1.1	2177	1	N60063	Human Factor VII	342	1.1	9080	1	V19580	Human factor III e
270	1.1	2189	1	N80944	Alpha-2 plasmin in	343	1.1	10660	1	Q84793	Spinocherebellar at
271	1.1	2209	1	T87427	Clone J5. Polynucl	344	1.1	11141	1	T15820	TBE virus strain N
272	1.1	2211	1	Q12190	Human sperm-specif	345	1.1	11141	1	T44469	Tick-borne encephal
273	1.1	2238	1	V24017	Human HBMBU14 codi	346	1.1	11517	1	Q26021	PSP6-SFV4 RNA tran
274	1.1	2243	1	Q13402	Human alpha-2 plas	347	1.1	12827	1	V09036	Equine arteritis v
275	1.1	2249	1	N92325	Human alpha-2-plas	348	1.1	15528	1	V09039	Equine arteritis v
276	1.1	2253	1	Q05998	Sequence encoding	349	1.1	29392	1	V15422	Mouse poly ig rece
277	1.1	2256	1	Q99002	Human acetylcholin	350	1.1	40	1	Q25031	Oligonucleotide sp
278	1.1	2280	1	V52604	Human metalloprote	351	1.1	44	1	V09273	Nucleotide sequenc
279	1.1	2335	1	Q80216	Human proNDP-alpha	352	1.1	48	1	T04081	Trypsin inhibitory
280	1.1	2346	1	V59693	Human secreted pro	353	1.1	49	1	V02146	Human secreted pro
281	1.1	2424	1	N91177	Sequence encoding	354	1.1	49	1	T88080	3' portion of cDNA
282	1.1	2432	1	X00650	Human secreted pro	355	1.1	50	1	Q66922	Poly-dA 50mer prob
283	1.1	2435	1	X00707	Human secreted pro	356	1.1	50	1	V37933	Primer of the spec
284	1.1	2438	1	N60065	Factor IX/Factor V	357	1.1	56	1	Q52732	Sequence of oligo
285	1.1	2445	1	N60111	Sequence encoding	358	1.1	56	1	Q52734	Sequence of oligo
286	1.1	2445	1	Q48393	Full-length foetal	359	1.1	61	1	Q78635	Human chondromodul
287	1.1	2541	1	X04343	Human secreted pro	360	1.1	61	1	V07498	Lelystad virus pri
288	1.1	2594	1	T02367	Melon aconitase cD	361	1.1	69	1	V02148	Human secreted pro
289	1.1	2605	1	V04699	Homo sapiens 20q13	362	1.1	69	1	T88081	3' portion of cDNA
290	1.1	2628	1	V01060	Human PKR gene. Sc	363	1.1	69	1	V21236	Homo sapiens clone
291	1.1	2634	1	X27062	S. tuberosum isoam	364	1.1	69	1	V21226	Homo sapiens clone
292	1.1	2643	1	T33855	Human urinary blad	365	1.1	70	1	V30923	Human secreted pro
293	1.1	2689	1	V43886	Human IgA nephropa	366	1.1	74	1	V99725	Human adult testis
294	1.1	2700	1	T97447	Pacific yew taxadi	367	1.1	77	1	V77194	Staphylococcus aur
295	1.1	2810	1	V17099	Human brain derive	368	1.1	79	1	V26706	Human novel secret
296	1.1	2880	1	X18193	hNBC3a coding sequ	369	1.1	79	1	V32414	Homo sapiens clone
297	1.1	3024	1	Q04049	Sequence of clone	370	1.1	80	1	V37197	Oligonucleotide se
298	1.1	3024	1	Q35260	ICAM-1 cDNA. Treat	371	1.1	80	1	V30930	Human secreted pro
299	1.1	3024	1	Q57397	Human ICAM-1. Inte	372	1.1	84	1	V04275	Secreted protein C
300	1.1	3024	1	T11672	Intracellular adhe	373	1.1	84	1	V09269	Nucleotide sequenc
301	1.1	3024	1	T47957	Human ICAM-1 (HL-6	374	1.1	85	1	T39470	Growth regulatory
302	1.1	3024	1	T62835	Human ICAM-1 cDNA.	375	1.1	85	1	V05720	Nucleotide sequenc
303	1.1	3024	1	V70950	Intracellular adhe	376	1.1	85	1	V37488	Human growth regul
304	1.1	3113	1	N80995	HL-cDNA clone enco	377	1.1	86	1	Q11760	Self-complementary
305	1.1	3114	1	V02800	Human RHAMM cDNA.	378	1.1	86	1	V11441	Human secreted pro
306	1.1	3176	1	T62124	Arabidopsis thalia	379	1.1	87	1	V05740	Nucleotide sequenc
307	1.1	3320	1	V56991	Maize starch phosph	380	1.1	90	1	Q11762	Self-complementary
308	1.1	3350	1	T18200	Infectious bursal	381	1.1	90	1	T47082	Synthetic DNA prob
309	1.1	3471	1	Q46673	p mySB cDNA clone.	382	1.1	90	1	T47081	Synthetic DNA prob
310	1.1	3547	1	N60846	Plasmid sequence e	383	1.1	90	1	V31773	Synthetic oligonuc
311	1.1	3572	1	V67162	Drosophila kinesin	384	1.1	90	1	V54077	Nucleotide sequenc
312	1.1	3592	1	Q53529	Human Mannose-Bind	385	1.1	92	1	V11595	Homo sapiens adult
313	1.1	3692	1	T72107	Rat semaphorin 2 g	386	1.1	92	1	V61480	Human secreted pro
314	1.1	3798	1	Q42424	ADA2 DNA. Transcri	387	1.1	94	1	V05728	Nucleotide sequenc
315	1.1	3912	1	X18192	hNBC3 coding seque	388	1.1	98	1	Q11764	Self-complementary
316	1.1	3933	1	T03885	Human mucosal lymph	389	1.1	98	1	X00174	Porcine reproducti
317	1.1	4180	1	N91773	Rat androgen recep	390	1.1	99	1	T91300	Human M97-2 secret
318	1.1	4364	1	Q63730	Protein tyrosine k	391	1.1	100	1	V68827	DNA molecule encod
319	1.1	4589	1	Q06068	p150.95 alpha subu	392	1.1	101	1	V00420	3' fragment of clo
320	1.1	4765	1	X09003	Mouse semaphorin r	393	1.1	112	1	V41441	Nucleotide sequenc
321	1.1	4780	1	X09004	Mouse semaphorin r	394	1.1	114	1	T39472	Growth regulatory
322	1.1	4985	1	X00463	Human type IX aden	395	1.1	114	1	V37490	Human growth regul
323	1.1	5555	1	Q64898	HER4 with alternat	396	1.1	117	1	V21223	Homo sapiens clone
324	1.1	5555	1	T18533	Receptor tyrosine	397	1.1	119	1	V44858	Clone Ad1_1_3' seq
325	1.1	5613	1	Q63241	Crucifer l-aminocyt	398	1.1	120	1	Q52429	Human subtelomeric

399	34	1.1	120	1	Q87011	Subtelomeric cDNA	472	34	1.1	570	1	T70048	Cotton fibre speci
400	34	1.1	127	1	T96338	Subtelomeric cDNA	473	34	1.1	572	1	X07182	Soybean threonine
401	34	1.1	120	1	T11604	Homo sapiens adult	474	34	1.1	577	1	V59806	Human secreted pro
402	34	1.1	140	1	T76781	Staphylococcus aur	475	34	1.1	579	1	X82799	ATG-1100 (allograf
403	34	1.1	141	1	T06012	Immunodominant fra	476	34	1.1	585	1	X37386	Human secreted pro
404	34	1.1	141	1	T34595	Probe for detectin	477	34	1.1	596	1	Q10488	Fusion protein enc
405	34	1.1	144	1	V21239	Homo sapiens clone	478	34	1.1	596	1	V00462	Human interleukin-
406	34	1.1	191	1	T13475	Capture probe for	479	34	1.1	600	1	Q72739	T. ovis vaccine ca
407	34	1.1	198	1	Q42784	Ligand-induced gen	480	34	1.1	600	1	T17934	Taenia ovis antige
408	34	1.1	201	1	V00418	3' fragment of clo	481	34	1.1	604	1	X37373	Human secreted pro
409	34	1.1	202	1	V41451	Nucleotide sequenc	c 482	34	1.1	609	1	T62617	Cotton fibre speci
410	34	1.1	208	1	V09116	3' nucleotide sequ	483	34	1.1	615	1	V02881	Human HMG1-C aberr
411	34	1.1	216	1	V09114	3' nucleotide sequ	484	34	1.1	618	1	V58362	Coding sequence fo
412	34	1.1	222	1	Q46071	Sequence downstrea	485	34	1.1	625	1	V40885	Coding sequence of
413	34	1.1	227	1	T89600	Hepatitis C virus	486	34	1.1	625	1	X19927	Lily LGC1 nucleoti
414	34	1.1	240	1	T76782	Staphylococcus aur	487	34	1.1	633	1	V09282	Nucleotide sequenc
415	34	1.1	249	1	T89608	Hepatitis C virus	488	34	1.1	634	1	X30403	DNA encoding a hum
416	34	1.1	257	1	T89599	Hepatitis C virus	489	34	1.1	644	1	X00684	Human secreted pro
417	34	1.1	260	1	T89601	Hepatitis C virus	490	34	1.1	645	1	X04348	Human secreted pro
418	34	1.1	270	1	T89602	Hepatitis C virus	491	34	1.1	660	1	X37406	Human secreted pro
419	34	1.1	286	1	X30353	DNA encoding a hum	492	34	1.1	667	1	Q61566	Human SOD. Recombi
420	34	1.1	289	1	Q53450	Sequence of the 3'	493	34	1.1	667	1	Q61573	Human SOD. Recombi
421	34	1.1	292	1	V21153	3' nucleotide port	494	34	1.1	667	1	Q61575	Hybrid human/plig S
422	34	1.1	296	1	X25862	Clone PEGen-44 seq	495	34	1.1	667	1	X04334	Human secreted pro
423	34	1.1	314	1	V09108	3' nucleotide sequ	496	34	1.1	688	1	Q61572	Human SOD. Recombi
424	34	1.1	321	1	T89598	Hepatitis C virus	497	34	1.1	688	1	V34216	Human secreted pro
425	34	1.1	334	1	X25857	Clone PEGen-24 seq	498	34	1.1	699	1	O58663	Human superoxide d
426	34	1.1	337	1	X25860	Clone PEGen-42 seq	499	34	1.1	700	1	Q22352	Pl4(?) allergen cl
427	34	1.1	340	1	Q99393	Rat allograft infl	500	34	1.1	700	1	Q89191	Birch pollen Pl4 a
428	34	1.1	340	1	V89136	EST clone BR309. N	501	34	1.1	700	1	V34301	Human secreted pro
429	34	1.1	342	1	Q59382	Human brain Expres	502	34	1.1	704	1	V35167	Thuga plicata diri
430	34	1.1	345	1	V26683	Human novel secret	503	34	1.1	704	1	V84441	Human secreted pro
431	34	1.1	347	1	Q27482	T23 cDNA. Stamen-s	504	34	1.1	719	1	V40501	Human secreted pro
432	34	1.1	350	1	Q99380	Rat allograft infl	505	34	1.1	722	1	X06786	Human adult uterus
433	34	1.1	356	1	T89597	Hepatitis C virus	c 506	34	1.1	724	1	V58515	3' fragment of pro
434	34	1.1	372	1	N80409	Cowpea trypsin inh	c 507	34	1.1	724	1	V61281	3' cDNA sequence o
435	34	1.1	373	1	X30409	DNA encoding a hum	508	34	1.1	728	1	T85876	Malassezia fungus
436	34	1.1	375	1	T84939	Human prostate pro	509	34	1.1	730	1	V00421	Clone B219.2. Isol
437	34	1.1	382	1	Q60162	Human brain Expres	510	34	1.1	737	1	V34270	Human secreted pro
438	34	1.1	399	1	T13044	Cotton fibre-speci	511	34	1.1	740	1	V23315	Plant CCR enzyme D
439	34	1.1	399	1	T62621	Cotton fibre cell-	512	34	1.1	741	1	V23878	Plant CCR enzyme D
440	34	1.1	399	1	T62620	Cotton fibre speci	513	34	1.1	742	1	V00437	Clone H905_107. Is
441	34	1.1	399	1	T70051	Cotton fibre speci	514	34	1.1	748	1	Q10376	Plasmid pMG4B12 us
442	34	1.1	413	1	X27371	Human secreted pro	515	34	1.1	755	1	X37467	Human secreted pro
443	34	1.1	423	1	X25073	Tobacco NADP-ferro	516	34	1.1	756	1	Q73729	Rape abscission/de
444	34	1.1	427	1	T85904	Malassezia fungus	517	34	1.1	760	1	V84625	Human secreted pro
445	34	1.1	429	1	V09097	Nucleotide sequenc	518	34	1.1	766	1	N40162	Sequence of prepro
446	34	1.1	429	1	V15712	Clone BV239_3 nucl	519	34	1.1	769	1	T79095	Full length human
447	34	1.1	443	1	N70773	Sequence encoding	520	34	1.1	772	1	V58363	Coding sequence fo
448	34	1.1	458	1	Q90525	Rat SIII 15 kDa su	521	34	1.1	773	1	X37388	Human secreted pro
449	34	1.1	458	1	T59893	Rat RNA polymerase	522	34	1.1	775	1	X26148	Human pan-s/tk-1A
450	34	1.1	464	1	V38682	Mus musculus SOCS1	523	34	1.1	776	1	X30318	DNA encoding a hum
451	34	1.1	467	1	V64423	Mouse developing l	524	34	1.1	778	1	X30336	DNA encoding a hum
452	34	1.1	472	1	V84530	Human secreted pro	525	34	1.1	779	1	Q79737	Flower style-speci
453	34	1.1	479	1	X32422	Homo sapiens clone	526	34	1.1	779	1	T64554	Tomato S-ribonucle
454	34	1.1	487	1	Q11686	Sequence encoding	527	34	1.1	785	1	V84473	Human secreted pro
455	34	1.1	487	1	X30156	Human secreted pro	528	34	1.1	787	1	T29008	Parietaria allerg
456	34	1.1	492	1	X22223	Human secreted pro	c 529	34	1.1	789	1	V58509	3' fragment of pro
457	34	1.1	513	1	X22214	Human secreted pro	c 530	34	1.1	789	1	V61275	3' cDNA sequence o
458	34	1.1	515	1	T13826	CHL1 neural cell a	531	34	1.1	791	1	X19493	Human secreted pro
459	34	1.1	524	1	V73005	Human adult testis	532	34	1.1	796	1	Q27948	CA455. Anther-spec
460	34	1.1	525	1	V41453	Nucleotide sequenc	533	34	1.1	801	1	V15416	Human chemokine MI
461	34	1.1	528	1	V02891	Human HMG1-C aberr	534	34	1.1	804	1	V59625	Human secreted pro
462	34	1.1	536	1	Q53448	Sequence encoding	535	34	1.1	809	1	V59739	Human secreted pro
463	34	1.1	546	1	X37513	Human secreted pro	536	34	1.1	812	1	T85877	Malassezia fungus
464	34	1.1	550	1	V41916	Nucleotide sequenc	537	34	1.1	819	1	V63198	cDNA from clone ga
465	34	1.1	553	1	X51720	DNA encoding a hum	538	34	1.1	824	1	X37387	Human secreted pro
466	34	1.1	556	1	V66764	Pathogen response	539	34	1.1	826	1	Q79736	Flower style-speci
467	34	1.1	562	1	V64661	Seq ID 29 from DE1	540	34	1.1	826	1	T64553	Tomato S-ribonucle
468	34	1.1	568	1	V53256	ETH precursor anti	541	34	1.1	826	1	X27402	Human secreted pro
469	34	1.1	568	1	V34054	Antisense full len	542	34	1.1	836	1	X04760	DNA encoding a coc
470	34	1.1	570	1	T13041	Cotton fibre-speci	543	34	1.1	859	1	T62944	Human eosinocyte C
471	34	1.1	570	1	T30258	Cotton fibre cell-	c 544	34	1.1	860	1	Q22007	Sequence encoding

545	34	1.1	861	1	Q99552	Thrombopoietin cod	618	1.1	1149	1	X27384	Human secreted pro
546	34	1.1	861	1	V40622	DN722.2 cDNA clone	619	1.1	1153	1	V59541	Human secreted pro
547	34	1.1	865	1	V48123	Nucleotide sequenc	620	1.1	1153	1	V55748	Human secreted pro
548	34	1.1	865	1	V84470	Human secreted pro	621	1.1	1154	1	T39048	cDNA encoding cell
549	34	1.1	865	1	X30358	DNA encoding a hum	622	1.1	1160	1	V40008	Human cell death a
550	34	1.1	873	1	V35164	Thuga plicata diri	623	1.1	1163	1	Q35198	ppcdHP71 coding se
551	34	1.1	879	1	Q71632	Apo-B RNA editing	624	1.1	1164	1	V07654	Nucleotide sequenc
552	34	1.1	880	1	T91714	Rat SmLIM/CRP2 cDN	625	1.1	1165	1	N60391	Sequence encoding
553	34	1.1	882	1	T72173	Alzheimer's diseas	626	1.1	1166	1	T60972	Mouse lipopolysacc
554	34	1.1	887	1	Q05160	Fragment of plasm	627	1.1	1169	1	V99983	3' untranslated re
555	34	1.1	887	1	Q06743	Insert from ptiGH6	628	1.1	1172	1	Q58669	Protein allergen o
556	34	1.1	889	1	V19304	Human p26 cDNA seq	c 629	1.1	1173	1	Q29636	Hepatitis C virus
557	34	1.1	890	1	X04376	Human secreted pro	c 630	1.1	1173	1	Q43894	NANB hepatitis vir
558	34	1.1	903	1	T47670	Human G protein ga	c 631	1.1	1173	1	Q43896	NANB hepatitis vir
559	34	1.1	906	1	Q80222	Human Nf-beta3 cl	c 632	1.1	1173	1	Q43897	NANB hepatitis vir
560	34	1.1	911	1	T27409	Human secreted pro	633	1.1	1174	1	T39050	cDNA encoding cell
561	34	1.1	916	1	T65651	cDNA encoding anti	634	1.1	1174	1	V39096	Monocomponent endo
562	34	1.1	916	1	V62000	R. prolixus NO-r p	c 635	1.1	1181	1	V81616	Cucumbr peroxidase
563	34	1.1	917	1	X27343	Human secreted pro	c 636	1.1	1181	1	V62867	Cucumbr peroxidase
564	34	1.1	921	1	V40521	Homo sapiens CH27-	637	1.1	1189	1	V41452	Nucleotide sequenc
565	34	1.1	934	1	X04394	Human secreted pro	638	1.1	1210	1	V84425	Human secreted pro
566	34	1.1	936	1	T69171	Typanosoma cruzi	639	1.1	1213	1	T61895	Human 14-3-3 beta
567	34	1.1	936	1	X51702	DNA encoding a hum	640	1.1	1215	1	T76885	Arabidopsis floral
568	34	1.1	938	1	V27142	Novel haemopoietin	641	1.1	1215	1	V58306	Arabidopsis thalia
569	34	1.1	941	1	N30062	Sequence of a modi	642	1.1	1218	1	T48447	Human TGF-beta-lik
570	34	1.1	943	1	V59535	Human secreted pro	643	1.1	1223	1	X25130	Soybean isoflavone
571	34	1.1	946	1	X37468	Human secreted pro	644	1.1	1227	1	T48446	Human TGF-beta-lik
572	34	1.1	953	1	T03910	Sequence encoding	645	1.1	1228	1	T48449	Human TGF-beta-lik
573	34	1.1	957	1	Q04441	Aequorin gene Bios	646	1.1	1229	1	Q38229	Sequence of rye gr
574	34	1.1	959	1	N81534	pA0440 aequorin ge	647	1.1	1229	1	Q85932	cDNA encoding Lol
575	34	1.1	961	1	T13151	Heparin-binding gr	648	1.1	1230	1	X22276	Human secreted pro
576	34	1.1	966	1	T87430	Clone H174-10. Pol	649	1.1	1230	1	X07565	Homo sapiens fetal
577	34	1.1	966	1	X04360	Human secreted pro	650	1.1	1232	1	V40526	Homo sapiens C2268
578	34	1.1	975	1	V61499	Bak binding protei	651	1.1	1239	1	N91576	Sequence of clone
579	34	1.1	985	1	N90529	Polydeoxyribonucle	652	1.1	1239	1	Q23000	Sequence encoding
580	34	1.1	987	1	X04331	Human secreted pro	653	1.1	1239	1	Q43439	Lol pi cDNA clone
581	34	1.1	990	1	X37471	Human secreted pro	654	1.1	1240	1	X06787	Human adult trache
582	34	1.1	997	1	X52274	Protein PRO244 cDN	655	1.1	1242	1	T13809	Murine Vrf186 cDNA
583	34	1.1	999	1	V80678	Human DNAX toll-li	656	1.1	1248	1	V11859	Homo sapiens Tub I
584	34	1.1	1001	1	X33229	Wheat viviparous l	657	1.1	1252	1	V32416	Homo sapiens clone
585	34	1.1	1034	1	Q12340	Caldesmon-like pol	658	1.1	1255	1	X37455	Human secreted pro
586	34	1.1	1015	1	X00713	Human secreted pro	659	1.1	1263	1	V99722	Human adult retina
587	34	1.1	1016	1	T29401	Lipase coding sequ	660	1.1	1265	1	V68898	DNA molecule encod
588	34	1.1	1020	1	X00681	Human secreted pro	661	1.1	1285	1	Q38028	CAD cDNA isolated
589	34	1.1	1025	1	X30319	DNA encoding a hum	662	1.1	1296	1	T04580	Human IFNAB-BPI. N
590	34	1.1	1032	1	X30155	Human secreted pro	663	1.1	1302	1	X30355	Nucleotide sequenc
591	34	1.1	1035	1	T60568	Cotton fibrous tis	664	1.1	1307	1	V40509	DNA encoding a hum
592	34	1.1	1041	1	Q54651	T84.12 light chain	665	1.1	1310	1	V59612	Human secreted pro
593	34	1.1	1041	1	Q54653	T84.12 L4-12-1 lig	666	1.1	1313	1	V59612	Human secreted pro
594	34	1.1	1045	1	Q79671	Sequence encoding	c 667	1.1	1315	1	T65022	Human nuclear prol
595	34	1.1	1047	1	Q30001	HCNP precursor gen	668	1.1	1315	1	V73311	Natural killer cel
596	34	1.1	1047	1	Q58685	Rat hippocampal ch	669	1.1	1316	1	T99543	Human chemokine re
597	34	1.1	1049	1	V43610	Human secreted pro	670	1.1	1322	1	V43616	Human secreted pro
598	34	1.1	1054	1	X00618	Human secreted pro	671	1.1	1323	1	T51738	Orange pectin meth
599	34	1.1	1065	1	X30407	DNA encoding a hum	672	1.1	1323	1	V64073	Orange pectin meth
600	34	1.1	1078	1	Q25785	Protein synthesis	673	1.1	1325	1	Q31873	Cyclin D1 gene. Re
601	34	1.1	1080	1	T04682	Human superoxide-d	674	1.1	1325	1	Q53197	Human cyclin D1. N
602	34	1.1	1083	1	V24126	Homo sapiens BE2 s	675	1.1	1325	1	V20473	Human cyclin D1 on
603	34	1.1	1090	1	X37521	Human secreted pro	676	1.1	1340	1	T85980	Human neuro-D gene
604	34	1.1	1096	1	X22275	Human secreted pro	677	1.1	1342	1	V62449	Human neurotensin
605	34	1.1	1113	1	Q29633	Hepatitis C virus	678	1.1	1348	1	Q78278	Human interleukin-
606	34	1.1	1114	1	V34315	Human secreted pro	679	1.1	1348	1	X24749	Human interleukin-
607	34	1.1	1118	1	Q20263	Deg-1 gene. bNA se	680	1.1	1364	1	T48099	Human interleukin-
608	34	1.1	1123	1	V34290	Human secreted pro	681	1.1	1364	1	V15228	Heterodimeric huma
609	34	1.1	1124	1	Q03669	Papaya ringspot vi	682	1.1	1364	1	V42536	Human interleukin-
610	34	1.1	1129	1	X27340	Human secreted pro	683	1.1	1367	1	X51727	DNA encoding a hum
611	34	1.1	1136	1	N70435	Sequence encoding	684	1.1	1370	1	X19886	Pancreas derived p
612	34	1.1	1139	1	V08813	Gene No. 3 encodin	685	1.1	1371	1	V17829	Homo sapiens pancr
613	34	1.1	1141	1	T13810	Murine Vrf167 cDNA	686	1.1	1371	1	X19885	Pancreas derived p
614	34	1.1	1142	1	X51744	DNA encoding a hum	687	1.1	1373	1	V52969	Human G1 protein i
615	34	1.1	1143	1	V59619	Human secreted pro	688	1.1	1376	1	Q56747	Ryegrass Lol pv al
616	34	1.1	1146	1	X07000	Xenopus Frazzled p	689	1.1	1376	1	X04325	Human secreted pro
617	34	1.1	1147	1	V69620	Human secreted pro	690	1.1	1378	1	V59706	Human secreted pro

691	34	1.1	1384	1	X37442	Human secreted pro	764	34	1.1	1606	1	Q20264	Mec-4 gene. DNA se
692	34	1.1	1393	1	X38027	CAD cDNA isolated	765	34	1.1	1607	1	Q38934	med-4 gene. Isolat
693	34	1.1	1393	1	Q66241	Alpha-galactosidas	766	34	1.1	1620	1	V43793	Rodent chemokine r
694	34	1.1	1393	1	Q77861	Human alpha-galact	767	34	1.1	1624	1	V17054	Glucosyl transfera
695	34	1.1	1394	1	Q05879	Placenta-specific	768	34	1.1	1627	1	V40486	Human secreted pro
696	34	1.1	1396	1	Q24371	Natural killer cel	769	34	1.1	1631	1	T72795	Beta-1 integrin mo
697	34	1.1	1400	1	Q06351	Recombinant urokin	770	34	1.1	1637	1	V69761	EDG-1-like G-prote
698	34	1.1	1400	1	T29191	Human urokinase pi	771	34	1.1	1640	1	Q50575	Asparaginylendopep
699	34	1.1	1400	1	X19823	Urokinase-type pro	772	34	1.1	1641	1	T15758	Human interleukin-
700	34	1.1	1404	1	V84622	Human secreted pro	773	34	1.1	1646	1	V69038	DNA molecule encod
701	34	1.1	1406	1	V61486	Human secreted pro	774	34	1.1	1653	1	X00699	Human secreted pro
702	34	1.1	1407	1	N60374	Sequence encoding	775	34	1.1	1654	1	T64943	CDNA encoding a no
703	34	1.1	1407	1	Q32515	Sequence encoding	776	34	1.1	1654	1	T93210	Human neurotactin
704	34	1.1	1407	1	Q71647	Human phospholipas	777	34	1.1	1659	1	T30360	Human p57 coding s
705	34	1.1	1407	1	T89675	Human phospholipas	778	34	1.1	1659	1	T31800	Maize adenylosuccin
706	34	1.1	1410	1	X15022	Human secreted pro	779	34	1.1	1668	1	X00664	Human secreted pro
707	34	1.1	1414	1	V33888	Hyocysamine 6 beta	780	34	1.1	1670	1	V26557	Human IP-10/Mig re
708	34	1.1	1420	1	T72167	H. tuberosus CYCD3;	781	34	1.1	1679	1	X00719	Human secreted pro
709	34	1.1	1420	1	X27321	Alzheimer's diseas	782	34	1.1	1682	1	Q91763	Human Rad51 gene,
710	34	1.1	1425	1	T35220	Human secreted pro	783	34	1.1	1686	1	V73012	Human adult brain
711	34	1.1	1434	1	X37452	Cytoplasmic antipr	784	34	1.1	1687	1	V01072	Bovine p58 gene. S
712	34	1.1	1434	1	T48444	Human secreted pro	785	34	1.1	1694	1	V33190	Secreted protein B
713	34	1.1	1441	1	X00631	Human TGF-beta-lik	786	34	1.1	1694	1	X07431	Homo sapiens secre
714	34	1.1	1448	1	X00631	Human secreted pro	787	34	1.1	1699	1	V34294	Human secreted pro
715	34	1.1	1453	1	Q50573	Asparaginylendopep	788	34	1.1	1700	1	V11858	Mus musculus Tub I
716	34	1.1	1454	1	V23913	Plant CAD enzyme D	789	34	1.1	1710	1	Q47966	Rape acyl-ACP thio
717	34	1.1	1461	1	T01471	Mammalian mast cel	790	34	1.1	1711	1	V43617	Human secreted pro
718	34	1.1	1469	1	V84120	Pentaclethra macro	791	34	1.1	1713	1	T85993	Maize 5-enolpyruv
719	34	1.1	1474	1	T90174	Oil seed rape cyst	792	34	1.1	1713	1	T86000	Maize 5-enolpyruv
720	34	1.1	1474	1	V23914	Plant CAD enzyme D	793	34	1.1	1713	1	V19277	Maize 5-enolpyruv
721	34	1.1	1483	1	T27147	Chalcone synthase	794	34	1.1	1717	1	V08170	MMP19 coding seque
722	34	1.1	1487	1	X00698	Human secreted pro	795	34	1.1	1718	1	X26151	3' CDNA sequence o
723	34	1.1	1490	1	Q75361	Gibberellin-20-ox	796	34	1.1	1721	1	T38129	Nuclear envelope-a
724	34	1.1	1493	1	Q04492	Sequence encoding	797	34	1.1	1721	1	X07566	Homo sapiens fetal
725	34	1.1	1493	1	X33994	Human osteopontin	798	34	1.1	1728	1	V38990	Human receptor for
726	34	1.1	1494	1	V41444	Nucleotide sequenc	799	34	1.1	1730	1	T61340	Batten disease gen
727	34	1.1	1502	1	T68721	Human mitogen-acti	800	34	1.1	1731	1	T61348	Batten disease gen
728	34	1.1	1502	1	V10099	Human MAP kinase p	801	34	1.1	1731	1	T61333	Batten disease gen
729	34	1.1	1503	1	X32253	Protein PRO272 cDN	802	34	1.1	1732	1	T61335	Batten disease gen
730	34	1.1	1505	1	T72177	Alzheimer's diseas	803	34	1.1	1732	1	T61336	Batten disease gen
731	34	1.1	1505	1	V40522	Homo sapiens CI542	804	34	1.1	1732	1	T61337	Batten disease gen
732	34	1.1	1507	1	X00656	Human secreted pro	805	34	1.1	1732	1	T61338	Batten disease gen
733	34	1.1	1508	1	V07665	Maize Bx1 cDNA (DI	806	34	1.1	1732	1	T61341	Batten disease gen
734	34	1.1	1509	1	V59679	Human secreted pro	807	34	1.1	1732	1	T61343	Batten disease gen
735	34	1.1	1510	1	Q03740	Human myb related	808	34	1.1	1732	1	T61344	Batten disease gen
736	34	1.1	1511	1	T29032	Endoglucanase (S8	809	34	1.1	1732	1	T61345	Batten disease gen
737	34	1.1	1514	1	Q05981	Insert from pTPRLI	810	34	1.1	1732	1	T61346	Batten disease gen
738	34	1.1	1515	1	V84440	Human secreted pro	811	34	1.1	1732	1	T61347	Batten disease gen
739	34	1.1	1536	1	Q94111	mML genomic DNA. T	812	34	1.1	1732	1	T61332	Batten disease gen
740	34	1.1	1537	1	V59804	Human secreted pro	813	34	1.1	1732	1	T61306	Batten disease cDN
741	34	1.1	1538	1	T63302	A. chrysoeum O-a	814	34	1.1	1733	1	T61339	Batten disease gen
742	34	1.1	1541	1	Q35199	ppcDHP23 coding se	815	34	1.1	1733	1	T61342	Batten disease gen
743	34	1.1	1549	1	X24921	Guman ras carboxy-	816	34	1.1	1734	1	T61334	Batten disease gen
744	34	1.1	1553	1	T68695	Strawberry inverta	817	34	1.1	1737	1	Q29506	New platelet facto
745	34	1.1	1553	1	V81285	Human RCE1 (hRCE1)	818	34	1.1	1737	1	Q80521	Human monocyte PP4
746	34	1.1	1558	1	Q10896	Encodes Xenopus Bo	819	34	1.1	1737	1	Q99007	Chemokine superfam
747	34	1.1	1558	1	T28255	Survival motor neu	820	34	1.1	1738	1	Q56241	Sequence encoding
748	34	1.1	1558	1	X22942	Human cathepsin Y	821	34	1.1	1741	1	T32830	Wheat adenylosuccin
749	34	1.1	1559	1	T02714	MART-1 melanoma an	822	34	1.1	1747	1	Q20260	ERK1. DNA encoding
750	34	1.1	1560	1	T18828	Human survival mot	823	34	1.1	1748	1	Q88327	Arabidopsis thalia
751	34	1.1	1562	1	Q42769	Ligand-induced gen	824	34	1.1	1749	1	V21800	Human spleen-deriv
752	34	1.1	1564	1	X22231	Human secreted pro	825	34	1.1	1752	1	V59556	Human secreted pro
753	34	1.1	1569	1	N90711	Recombinant activa	826	34	1.1	1753	1	T33235	Soluble starch syn
754	34	1.1	1575	1	N90112	Anticoagulant PP4	827	34	1.1	1758	1	V14145	Nucleotide sequenc
755	34	1.1	1577	1	T90169	Oil seed rape cyst	828	34	1.1	1761	1	X20412	Human secreted pro
756	34	1.1	1579	1	V64590	Tobacco chlI DNA.	829	34	1.1	1769	1	T85973	Human NB Phox 3' u
757	34	1.1	1582	1	T18831	Human survival mot	830	34	1.1	1770	1	X52242	Protein PRO233 cDN
758	34	1.1	1582	1	T28259	Survival motor neu	831	34	1.1	1772	1	V63174	CDNA from clone dd
759	34	1.1	1583	1	V58758	Human secreted pro	832	34	1.1	1776	1	T91744	Tobacco calcium/ca
760	34	1.1	1586	1	Q47967	Rape acyl-ACP thio	833	34	1.1	1780	1	V04269	Rat FRAG1 coding s
761	34	1.1	1587	1	X06840	Arabidopsis branch	834	34	1.1	1793	1	Q15139	Clone pTACC6 encod
762	34	1.1	1594	1	V62755	Human secreted pro	835	34	1.1	1801	1	N70524	Sequence of cDNA 1
763	34	1.1	1604	1	V04252	Wild tomato acylr	836	34	1.1	1813	1	Q88760	Human ubiquitous n

837	1.1	1817	1	V04064	Human RNA-binding	910	1.1	2120	1	X33812	Coding sequence fo
838	34	1821	1	V59590	Human secreted pro	911	34	2123	1	T59701	PTH-like peptide D
839	34	1824	1	Q47842	Flavonoid-3',5'-hy	912	34	2143	1	Q90242	Full length Sorghu
840	34	1829	1	V26298	Phosphate starvati	913	34	2153	1	V84565	Human secreted pro
841	34	1830	1	X00612	Human secreted pro	914	34	2154	1	V57200	Human secreted pro
842	34	1834	1	Q04690	Encodes Mammalian	915	34	2159	1	Q67223	Mouse p55Nuc. Ant1
843	34	1835	1	X27245	Human CIAR1 coding	916	34	2196	1	V59663	Human secreted pro
844	34	1836	1	V33461	Human T1-receptor	917	34	2200	1	Q11010	Fibulin C. Purifie
845	34	1837	1	T33852	Rat vas deferens p	918	34	2205	1	V04680	Human presenilin 1
846	34	1841	1	Q91311	Human Fas-L gene.	919	34	2214	1	V40524	Homo sapiens COL22
847	34	1841	1	X21574	Human FasL protein	920	34	2218	1	V84503	Human secreted pro
848	34	1841	1	X24879	Human Fas ligand (921	34	2218	1	X02887	Rat mACHR-6 cDNA.
849	34	1841	1	V24018	Human BHF1 coding	922	34	2223	1	Q39051	Tomato acid invert
850	34	1848	1	V84588	Human secreted pro	923	34	2229	1	Q23462	Human B-raf protei
851	34	1849	1	T74042	Soybean thiol prot	924	34	2230	1	V15704	Tomato ACC synthas
852	34	1858	1	T72175	Alzheimer's diseas	925	34	2233	1	V10120	Human retinoid rec
853	34	1863	1	N60399	Sequence of the pr	926	34	2239	1	V01527	Wheat soluble star
854	34	1864	1	V59630	Human secreted pro	927	34	2255	1	N60228	Sequence of the br
855	34	1866	1	V84466	Human secreted pro	928	34	2255	1	Q84698	Human IFN-gamma ac
856	34	1868	1	T73502	Poplar 1-aminocycl	929	34	2264	1	X20496	Human secreted pro
857	34	1875	1	V06408	Epidermal growth f	930	34	2265	1	X20435	Human secreted pro
858	34	1875	1	T48669	Human EDG-2 recept	931	34	2271	1	V84632	Human secreted pro
859	34	1875	1	V33033	Dirofilaria immiti	932	34	2276	1	V84583	Human secreted pro
860	34	1875	1	V33017	Dirofilaria immiti	933	34	2294	1	V79584	Rat organic anion
861	34	1877	1	V46370	Nucleic acid encod	934	34	2298	1	X06781	Human testis secre
862	34	1883	1	V34217	Human secreted pro	935	34	2304	1	Q14176	Clone pXR2C8 encod
863	34	1890	1	V34217	Human secreted pro	936	34	2304	1	Q55374	PAR2C8 DNA. Multim
864	34	1893	1	V09852	Pig plasma membran	937	34	2304	1	T76787	Insect XR2C recept
865	34	1898	1	T79634	DNA encoding human	938	34	2304	1	T89959	D. melanogaster XR
866	34	1907	1	V84508	Human secreted pro	939	34	2307	1	V44318	Maize starch-associ
867	34	1908	1	T73002	Human adult brain	940	34	2312	1	V21646	T. gondii chorisma
868	34	1910	1	Q50579	Asparaginylendopep	941	34	2323	1	V59524	Human secreted pro
869	34	1910	1	V27038	Human napsin B cDN	942	34	2336	1	X02889	Human secreted pro
870	34	1922	1	N70525	Sequence of cDNA i	943	34	2339	1	Q63892	Rat mACHR-6 cDNA f
871	34	1923	1	N91039	DNA sequence of ri	944	34	2339	1	T01339	Tomato acid invert
872	34	1925	1	T34271	Human secreted pro	945	34	2345	1	X18951	Human PIGR-1 encod
873	34	1931	1	Q12417	bymf gene encoding	946	34	2361	1	X04316	Human secreted pro
874	34	1934	1	T13229	OR-1 orphan recept	947	34	2377	1	Q20360	Human pro-urokinas
875	34	1946	1	V40886	Coding sequence of	948	34	2378	1	T84975	Rat glial cell der
876	34	1951	1	X37414	Human secreted pro	949	34	2400	1	Q05055	Placenta-specific
877	34	1958	1	V10328	Rat Galk2 receptor	950	34	2400	1	T93364	Nicotiana plumbagi
878	34	1961	1	T70155	BT20 protein encod	951	34	2405	1	Q70827	Protein kinase (CK
879	34	1964	1	N81586	CBT-20 cDNA encodi	952	34	2405	1	Q92962	Human HRR25-like c
880	34	1964	1	N92781	cDNA sequence enco	953	34	2424	1	V59725	Human secreted pro
881	34	1964	1	Q54354	Carcinoembryonic a	954	34	2427	1	Q04107	Human pro-urokinas
882	34	1964	1	T45064	Carcinoembryonic a	955	34	2430	1	Q83161	DNA encoding malic
883	34	1969	1	Q47671	Sequence of the hu	956	34	2431	1	Q31546	Human heregulin-be
884	34	1971	1	X08428	Acidic leucine ami	957	34	2447	1	V54587	Human secretory pr
885	34	1972	1	Q95309	MPD-1 cDNA. Prepar	958	34	2460	1	V44303	Human secreted pro
886	34	1973	1	N91045	Gene encoding bovi	959	34	2468	1	V40487	Human secreted pro
887	34	1973	1	Q55770	Murine programmed	960	34	2469	1	V82437	Rat 25-hydroxyvita
888	34	1987	1	X27344	Human secreted pro	961	34	2481	1	T87940	Rat neurotrophin
889	34	1994	1	T09085	Achatina fulica cD	962	34	2483	1	V84468	Human secreted pro
890	34	2000	1	Q74367	Bovine beta3 adren	963	34	2487	1	X33813	Coding sequence fo
891	34	2002	1	V43605	Human secreted pro	964	34	2510	1	X21931	Human B-raf kinase
892	34	2002	1	V63197	cDNA from clone fy	965	34	2519	1	V54124	Human membrane pro
893	34	2006	1	Q29607	Human kidney PTH/P	966	34	2524	1	T18673	Human trophoblast
894	34	2007	1	V40754	C. felis esterase,	967	34	2549	1	N91467	Sequence of human
895	34	2007	1	V40755	C. felis esterase,	968	34	2550	1	V19334	Rat brain ICAM-4 p
896	34	2010	1	T15948	Human kidney PTH/P	969	34	2550	1	V11663	Rat ICAM-4 clone 7
897	34	2010	1	V08391	Human Parathyroid	970	34	2550	1	V34661	Rat brain cDNA clo
898	34	2017	1	V08830	Gene No. 20 encodi	971	34	2550	1	V08980	Rat ICAM-4 coding
899	34	2026	1	V45313	Human secreted pro	972	34	2587	1	Q11148	GTP-gene from Chin
900	34	2032	1	T42073	Human nuclear prol	973	34	2589	1	T61590	Human c-IAP1. Nucl
901	34	2035	1	N70687	DNA encoding human	974	34	2601	1	T61591	Human c-IAP2. Nucl
902	34	2051	1	Q14944	Rhizopus PGK1 prom	975	34	2604	1	T73595	Rat neurotrophin
903	34	2057	1	T63123	Human metalloprote	976	34	2619	1	T61406	MACH isoform alpha
904	34	2060	1	V44862	Clone BV291_3 codi	977	34	2619	1	V74173	Human huBUB3 gene.
905	34	2095	1	X27331	Human secreted pro	978	34	2628	1	T01014	Lupin exo-(1-4)bet
906	34	2096	1	T36141	Human L-asparagina	979	34	2634	1	V59067	Bax inhibitor BI-1
907	34	2096	1	V15795	cDNA for wild type	980	34	2639	1	V52934	Pig transforming g
908	34	2104	1	V40506	Human secreted pro	981	34	2643	1	V99984	Nucleotide sequenc
909	34	2117	1	T61258	560E1 cDNA encodin	982	34	2662	1	Q37757	PVMO51. DNA sequen

983 34 1.1 2665 1 V32592 Schwannomin-bindin
 984 34 1.1 2669 1 Q56925 Pig TGF-beta-3. Nu
 985 34 1.1 2671 1 Q03303 Entire porcine tra
 986 34 1.1 2676 1 Q02819 cDNA sequence enco
 987 34 1.1 2682 1 V63195 CDNA from clone fg
 988 34 1.1 2696 1 N60909 Plasmid pAU157 ins
 989 34 1.1 2697 1 N60889 Plasmid pAU157 seq
 990 34 1.1 2697 1 N60864 Sequence of plasm
 991 34 1.1 2744 1 V84799 Nucleotide sequenc
 992 34 1.1 2760 1 Q85425 T2 DNA-polymerase
 993 34 1.1 2760 1 Q85426 T4 DNA-polymerase
 994 34 1.1 2764 1 V63190 CDNA from clone dh
 995 34 1.1 2773 1 X07428 Homo sapiens secre
 996 34 1.1 2788 1 V55745 Human secreted pro
 997 34 1.1 2790 1 V32998 Human concentrativ
 998 34 1.1 2822 1 X52222 Protein PRO228 cDN
 999 34 1.1 2823 1 T32233 Natural killer lyl
 1000 34 1.1 2836 1 V40744 C. felis esterase,

ALIGNMENTS

RESULT 1
 T99077/c
 ID T99077 standard; DNA; 45 BP.
 AC T99077;
 DT 24-MAR-1998 (first entry)
 DE Probe Ena-A7 for alphaENAC coding sequence.
 KW Alpha epithelial sodium channel; alphaENACA; binding assay;
 KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
 KW salt substitute; salty taste blocker; probe; ss.
 OS Synthetic.
 OS Rattus rattus.
 PN US5693756-A.
 PD 02-DEC-1997.
 PF 23-JAN-1995; 376362.
 PR 23-JAN-1995; US-376362.
 PR 28-FEB-1994; US-202654.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI Blackshaw S, Li X, Snyder SH;
 DT WPI: 98-031814/03.
 PT Alternatively spliced epithelial sodium channel alpha subunit
 PT proteins - useful in screening assays for salty taste enhancers or
 PT blockers
 PT Disclosure; Column 6; 33pp; English.
 CC This sequence represents a probe for the coding sequence for the alpha
 CC epithelial sodium channel a (alphaENACA). AlphaENACA (see W34529) and
 CC alphaENACb (see W34530) represent the sequences of the invention. The two
 CC sodium channels are alternatively spliced forms of the
 CC amiloride-sensitive salt channel alpha subunit and can be used in
 CC membrane-transport or binding assays to identify substances that enhance
 CC or block perception of a salty taste. Enhancers could be used as salt
 CC substitutes and blockers could be used to mask salty tastes in foods and
 CC pharmaceuticals.
 SQ Sequence 45 BP; 10 A; 10 C; 12 G; 13 T;

Query Match 1.4%; Score 45; DB 1; Length 45;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1134 caacaatggttttcctgcactgcacagacagacgaatgactt 1178
 |
 45 CAACAATGGTTGTCCTGCACACTGCGACAGACGACGAATGACTT 1

RESULT 2
 T99078/c
 ID T99078 standard; DNA; 45 BP.
 AC T99078;
 DT 24-MAR-1998 (first entry)
 DE Probe Ena-A9 for alphaENAC coding sequence.

KW Alpha epithelial sodium channel; alphaENACA; binding assay;
 KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
 KW salt substitute; salty taste blocker; probe; ss.
 OS Synthetic.
 OS Rattus rattus.
 PN US5693756-A.
 PD 02-DEC-1997.
 PF 23-JAN-1995; 376362.
 PR 23-JAN-1995; US-376362.
 PR 28-FEB-1994; US-202654.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI Blackshaw S, Li X, Snyder SH;
 DT WPI: 98-031814/03.
 PT Alternatively spliced epithelial sodium channel alpha subunit
 PT proteins - useful in screening assays for salty taste enhancers or
 PT blockers
 PT Disclosure; Column 7; 33pp; English.
 CC This sequence represents a probe for the coding sequence for the alpha
 CC epithelial sodium channel a (alphaENACA). AlphaENACA (see W34529) and
 CC alphaENACb (see W34530) represent the sequences of the invention. The two
 CC sodium channels are alternatively spliced forms of the
 CC amiloride-sensitive salt channel alpha subunit and can be used in
 CC membrane-transport or binding assays to identify substances that enhance
 CC or block perception of a salty taste. Enhancers could be used as salt
 CC substitutes and blockers could be used to mask salty tastes in foods and
 CC pharmaceuticals.
 SQ Sequence 45 BP; 11 A; 11 C; 10 G; 13 T;

Query Match 1.4%; Score 45; DB 1; Length 45;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1498 gactaccgaagcagagctctctgggctattgctattataaactg 1542
 |
 45 GACTACCGAAGCAGAGCTCTCTGGGCTATTGCTATTATAAAGT 1

RESULT 3
 T99072
 ID T99072 standard; DNA; 75 BP.

AC T99072;
 DT 24-MAR-1998 (first entry)
 DE Alpha epithelial sodium channel a coding sequence.
 KW Alpha epithelial sodium channel; alphaENACA; binding assay;
 KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
 KW salt substitute; salty taste blocker; ss.
 OS Rattus rattus.
 PN US5693756-A.
 PD 02-DEC-1997.
 PF 23-JAN-1995; 376362.
 PR 23-JAN-1995; US-376362.
 PR 28-FEB-1994; US-202654.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI Blackshaw S, Li X, Snyder SH;
 DT WPI: 98-031814/03.
 DR P-PSDB; W34529.
 PT Alternatively spliced epithelial sodium channel alpha subunit
 PT proteins - useful in screening assays for salty taste enhancers or
 PT blockers
 PS Disclosure; Column 21-22; 33pp; English.
 CC This sequence represents the coding sequence for the alpha epithelial
 CC sodium channel a (alphaENACA). AlphaENACA and alphaENACb (see W34530)
 CC represent the sequences of the invention. The two sodium channels are
 CC alternatively spliced forms of the amiloride-sensitive salt channel alpha
 CC subunit and can be used in membrane-transport or binding assays to
 CC identify substances that enhance or block perception of a salty taste.
 CC Enhancers could be used as salt substitutes and blockers could be used to
 CC mask salty tastes in foods and pharmaceuticals.
 SQ Sequence 75 BP; 11 A; 22 C; 22 G; 20 T;

Query Match 1.2%; Score 38; DB 1; Length 75;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 tctccaagtcgtcggaagcctttagtgcatacaactac 1614
|||||
DB 38 TCTCCAAAGTGTGGAGAGCCTGTAGTGTATCAACTAC 75

RESULT 4

Q03293
ID 003293 standard; DNA; 960 BP.
AC Q03293;
DT 03-AUG-1990 (first entry)
DE Recombinant DNA encoding human prolactin (HP)
KW Human prolactin (HP); plasmid pRP100; plasmid pDR720; plasmid pLP100.
OS Homo sapiens.
PN J02000445-A.
PD 05-JAN-1990.
PF 25-DEC-1987; 331244.
PR 25-DEC-1987; JP-331244, JP-315317.
PA (SHIK-) Shikishima Boseki K.

PI WPI; 90-047987/07.
DR P-PSDB; R05231.
PT In which promoter, Shine-Dalgarno sequence and translation
PT Initiation codon are integrated
PS Disclosure; Fig 1; 15pp; Japanese.
CC Also new are bacteria (E. coli) expressing HP which contain it and the
CC prodn. of HP by their culture. A promoter, Shine-Dalgarno (SD)
CC sequence and translation initiation codon (TIC) are inserted in
CC order, upstream of the HP gene. The HP gene is isolated from pLP100, the
CC promoter, SD sequence and TIC are obtd. from pDR720, and these are
CC ligated with a synthetic DNA linker to produce plasmid pRP100. It allows
CC large amts. of HP to be produced recombinantly.
SQ Sequence 960 BP; 287 A; 246 C; 209 G; 218 T;

Query Match 1.2%; Score 38; DB 1; Length 960;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 898 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 935

RESULT 5

V29126/c
ID V29126 standard; cDNA; 1276 BP.
AC V29126;
DT 21-AUG-1998 (first entry)
DE HPRP-1 coding sequence.
KW HPRP-1; human; disease associated protein tyrosine phosphatase; HPRP-2;
KW inflammatory cell; cancer; rheumatoid arthritis; osteoarthritis;
KW carcinoma; therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 604..1116
FT /tag=a
FT /product= HPRP-1

PN WO9814596-A2.
PD 09-APR-1998.
PE 02-OCT-1997; U17890.
PR 03-OCT-1996; US-725532.
PA (INCY-) INCYTE PHARM INC.
PI G011 SK; 98-240095/21.
DR WPI; 98-240095/21.
DR P-PSDB; W37913.
PT develop products for treating e.g. inflammatory conditions such as
PT rheumatoid arthritis or osteoarthritis, or cancers
PS Claim 3; Fig 1; 58pp; English.

CC This sequence encodes the human disease associated protein tyrosine
CC phosphatase protein, designated HPRP-1, of the invention. HPRP-1 and
CC HPRP-2 are associated with inflammatory cells and various cancers. They
CC can be used to develop products for treating inflammatory conditions such
CC as rheumatoid arthritis and osteoarthritis and carcinomas of the
CC intestine, bladder, prostate, breast and brain. The products can also be
CC used for detection, diagnosis and drug screening.
SQ Sequence 1276 BP; 407 A; 223 C; 225 G; 408 T;

Query Match 1.2%; Score 38; DB 1; Length 1276;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 41 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 4

RESULT 6

N60801
ID N60801 standard; DNA; 3557 BP.
AC N60801;
DT 24-OCT-1991 (first entry)
DE Human pro-growth hormone.
KW Pituitary gland; E.coli; ds.
OS Homo sapiens.
PN J61202689-A.
PD 08-SEP-1986.
PF 04-MAR-1985; 042404.
PR 04-MAR-1985; JP-042404.
PA (NAKA/) NAKAJIMA K.
DR WPI; 86-276387/42.
PT Plasmid for amplification of human growth hormone gene - derived
PT from human pituitary gland tissue.
PS Disclosure; Table 1-9; 9pp; Japanese.
CC Sequence is derived from human pituitary gland tissue, and allows
CC pro-growth hormone to be produced by a circular plasmid carrying an
CC amp resistance gene, from a transformed E.coli expression system in
CC an ampicillin containing medium.
SQ Sequence 3557 BP; 927 A; 902 C; 897 G; 831 T;

Query Match 1.2%; Score 38; DB 1; Length 3557;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 1149 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 1186

RESULT 7

N60847
ID N60847 standard; DNA; 3621 BP.
AC N60847;
DT 24-OCT-1991 (first entry)
DE Human pre-prolactin gene.
KW Pituitary gland; E.coli; ds.
OS Homo sapiens.
PN J61202690-A.
PD 08-SEP-1986.
PF 04-MAR-1985; 042405.
PR 04-MAR-1985; JP-042405.
PA (NAKA/) NAKAJIMA K.
DR WPI; 86-276388/42.

PT Plasmid for amplification of the human prolactin gene - derived
PT from human pituitary tissue and cultivated in ampicillin contg.
PT medium to obtain a pBR 322 ampicillin resistant gene.
PS Disclosure; Table 1-9; 9pp; Japanese.
CC Sequence is derived from human pituitary gland tissue, and allows
CC pre-prolactin to be produced by a circular plasmid carrying an amp

```
CC resistance gene, from a transformed E.coli expression system in an
CC amplicillin containing medium.
CC See also J61202689.
SQ Sequence 3621 BP; 969 A; 882 C; 894 G; 876 T;

Query Match 1.2%; Score 38; DB 1; Length 3621;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 1214 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 1251

RESULT 8
ID T74484 standard; cDNA; 495 BP.
AC T74484;
DT 01-NOV-1997 (first entry)
DE pilosulin 2 precursor cDNA clone Myr p2.
KW pilosulin 2; Australian jumper ant; venom; allergen; hypotensive;
KM blood pressure; cell growth; ss.
OS Myrmecia pilosula.
FH Key Location/Qualifiers
FT mat_peptide 44..271
FT signal_peptide 44..121
FT FT /tag- a
FT FT /tag- b
FT mat_peptide 188..268
FT FT /tag- C
FT FT /product- pilosulin A
PN MO9713854-A1.
PD 17-APR-1997.
PF 09-OCT-1996; AU0632.
PR 09-OCT-1995; AT-005839.
PA (NSYP-) NORTHERN SYDNEY AREA HEALTH SERVICE.
PI Baldo B. Donovan G;
DR WPI: 97-235890/21.
P-PSDB: W21784.
PT New isolated pilosulin 2 polypeptide - obtained from Myrmecia
PT pilosula ant venom, used for reducing blood pressure in animals or
PT for inhibiting cell growth.
PS Claim 1; Fig 1; 27pp; English.
CC This sequence is an isolated cDNA clone, designated Myr p2, which
CC includes a sequence encoding a polypeptide that has the biological
CC activity of pilosulin 2 having the amino acid sequence from
CC residues 49-75 of a 75-amino acid sequence (W21784). It was
CC isolated from a cDNA library prepared from the abdomens of Myrmecia
CC pilosula ants by immunoscreening with sera from venom-allergic
CC individuals. The isolated sequence can be incorporated into an
CC expression vector and used to produce pilosulin 2 in transformed
CC host cells. Pilosulin 2 and its analogues or fragments can be used
CC in claimed methods for reducing the blood pressure of an animal,
CC and for inhibiting the growth of (killing) cells.
SQ Sequence 495 BP; 201 A; 68 C; 99 G; 127 T;

Query Match 1.2%; Score 37; DB 1; Length 495;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 443 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 479

RESULT 9
ID V73439 standard; cDNA; 606 BP.
AC V73439;
DT 26-APR-1999 (first entry)
DE Flea saliva protein nucleic acid nfspl-606.
SQ Sequence 606 BP; 276 A; 80 C; 77 G; 173 T;
```

```
KW Flea saliva protein; nfspl-606; allergic dermatitis; allergen;
KM allergy; therapy; diagnosis; vaccine; ectoparasite; ss.
OS Ctenocephalides sp.
FH Key Location/Qualifiers
FT CDS 43..384
FT sig_peptide 43..93
FT FT /tag- a
FT mat_peptide 94..381
FT FT /tag- b
FT FT /tag- C
PN WO9845408-A2.
PD 15-OCT-1998.
PF 15-OCT-1997; U18669.
PR 10-APR-1997; WO-U05959.
PA (HESK-) HESKA CORP.
PI Weber ER;
DR WPI: 98-594480/50.
P-PSDB: W82386.
PT New isolated ectoparasite saliva genes - used to develop products
PT for the diagnosis, prevention, treatment and determining
PT susceptibility to allergic dermatitis
PS Claim 1; Page 142-143; 172pp; English.
CC This is the nucleotide sequence of nfspl-606, which encodes a
CC claimed full-length flea saliva protein, termed pfsppl-113 (see
CC W82386). nfspl-606 was isolated from a whole flea cDNA
CC library using partial clone nfspl-420 (see V73439) as probe.
CC Its nucleotide sequence shows no significant homology to known
CC sequences. Mature fsppl protein (see W82387), and a nucleic acid
CC sequence encoding it (see V73443) are also claimed. The invention
CC is directed to methods for isolating ectoparasite saliva proteins
CC (ESPs). It provides ESps (claimed, see W82382-93), nucleic acid
CC molecules encoding them, methods for their recombinant production,
CC therapeutic compositions for treating allergic dermatitis that
CC comprise at least one ESP, as well as assay kits for testing if an
CC animal has, or is susceptible to, allergic dermatitis, and a method
CC of desensitising a host animal to allergic dermatitis using ESps.
CC The ESps can also be used for the production of antibodies useful
CC in diagnosis or in vaccines for passive immunisation against
CC allergic dermatitis.
SQ Sequence 606 BP; 276 A; 80 C; 77 G; 173 T;

Query Match 1.2%; Score 37; DB 1; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 563 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 599

RESULT 10
ID V73440/C
AC V73440;
DT 26-APR-1999 (first entry)
DE Flea saliva protein nucleic acid nfspl-606 (complement).
KM Flea saliva protein; nfspl-606; allergic dermatitis; allergen;
KM allergy; therapy; diagnosis; vaccine; ectoparasite; ss.
OS Ctenocephalides sp.
PN WO9845408-A2.
PD 15-OCT-1998.
PF 15-OCT-1997; U18669.
PR 10-APR-1997; WO-U05959.
PA (HESK-) HESKA CORP.
PI Weber ER;
DR WPI: 98-594480/50.
PT New isolated ectoparasite saliva genes - used to develop products
PT for the diagnosis, prevention, treatment and determining
PT susceptibility to allergic dermatitis
PS Claim 1; Page 143; 172pp; English.
CC This is the nucleotide sequence of nucleic acid nfspl-606
CC complementary strand. The coding strand (see V73439) encodes a
```

CC claimed full-length flea saliva protein, termed Pfspj1-113 (see
CC W82386). nfspj1-606 was isolated from a whole flea cDNA
CC library using partial clone nfspj-420 (see V73419) as probe.
CC Its nucleotide sequence shows no significant homology to known
CC sequences. Mature fspj1 protein (see W82387), and a nucleic acid
CC sequence encoding it (see V73443) are also claimed. The invention
CC is directed to methods for isolating ectoparasite saliva proteins
CC (ESPs). It provides ESps (claimed, see W82382-93), nucleic acid
CC molecules encoding them, methods for their recombinant production,
CC therapeutic compositions for treating allergic dermatitis that
CC comprise at least one ESP, as well as assay kits for testing if an
CC animal has, or is susceptible to, allergic dermatitis, and a method
CC of desensitising a host animal to allergic dermatitis using ESps.
CC The ESps can also be used for the production of antibodies useful
CC in diagnosis or in vaccines for passive immunisation against
CC allergic dermatitis.
SQ Sequence 606 BP; 173 A; 77 C; 80 G; 276 T;

Query Match 1.2%; Score 37; DB 1; length 606;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 44 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 8

RESULT 11
V59566
ID V59566 standard; DNA: 732 BP.
AC 06-JAN-1999 (first entry)
DT Human secreted protein gene 56 clone HSAXS65.
DE Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; lissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; de;
KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9839448-A2.
PD 11-SEP-1998.
PE 06-MAR-1998; U04493.
PR 02-OCT-1997; US-061060.
PR 07-MAR-1997; US-038621.
PR 07-MAR-1997; US-040161.
PR 07-MAR-1997; US-040162.
PR 07-MAR-1997; US-040163.
PR 07-MAR-1997; US-040333.
PR 07-MAR-1997; US-040334.
PR 07-MAR-1997; US-040336.
PR 07-MAR-1997; US-040626.
PR 11-APR-1997; US-043311.
PR 11-APR-1997; US-043312.
PR 11-APR-1997; US-043313.
PR 11-APR-1997; US-043314.
PR 11-APR-1997; US-043568.
PR 11-APR-1997; US-043569.
PR 11-APR-1997; US-043576.
PR 11-APR-1997; US-043580.
PR 11-APR-1997; US-043669.
PR 11-APR-1997; US-043670.
PR 11-APR-1997; US-043671.
PR 11-APR-1997; US-043672.
PR 11-APR-1997; US-043674.
PR 23-MAY-1997; US-047492.
PR 23-MAY-1997; US-047500.
PR 23-MAY-1997; US-047501.
PR 23-MAY-1997; US-047502.

PR 23-MAY-1997; US-047503.
PR 23-MAY-1997; US-047581.
PR 23-MAY-1997; US-047582.
PR 23-MAY-1997; US-047583.
PR 23-MAY-1997; US-047584.
PR 23-MAY-1997; US-047585.
PR 23-MAY-1997; US-047586.
PR 23-MAY-1997; US-047587.
PR 23-MAY-1997; US-047588.
PR 23-MAY-1997; US-047589.
PR 23-MAY-1997; US-047590.
PR 23-MAY-1997; US-047592.
PR 23-MAY-1997; US-047593.
PR 23-MAY-1997; US-047594.
PR 23-MAY-1997; US-047595.
PR 23-MAY-1997; US-047596.
PR 23-MAY-1997; US-047597.
PR 23-MAY-1997; US-047598.
PR 23-MAY-1997; US-047599.
PR 23-MAY-1997; US-047600.
PR 23-MAY-1997; US-047601.
PR 23-MAY-1997; US-047612.
PR 23-MAY-1997; US-047613.
PR 23-MAY-1997; US-047614.
PR 23-MAY-1997; US-047615.
PR 23-MAY-1997; US-047617.
PR 23-MAY-1997; US-047618.
PR 23-MAY-1997; US-047632.
PR 23-MAY-1997; US-047633.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048974.
PR 13-JUN-1997; US-049610.
PR 08-JUL-1997; US-051926.
PR 16-JUL-1997; US-052874.
PR 18-AUG-1997; US-055724.
PR 22-AUG-1997; US-056630.
PR 22-AUG-1997; US-056631.
PR 22-AUG-1997; US-056632.
PR 22-AUG-1997; US-056636.
PR 22-AUG-1997; US-056637.
PR 22-AUG-1997; US-056638.
PR 22-AUG-1997; US-056679.
PR 22-AUG-1997; US-056880.
PR 22-AUG-1997; US-056881.
PR 22-AUG-1997; US-056882.
PR 22-AUG-1997; US-056884.
PR 22-AUG-1997; US-056886.
PR 22-AUG-1997; US-056887.
PR 22-AUG-1997; US-056888.
PR 22-AUG-1997; US-056889.
PR 22-AUG-1997; US-056892.
PR 22-AUG-1997; US-056893.
PR 22-AUG-1997; US-056894.
PR 22-AUG-1997; US-056903.
PR 22-AUG-1997; US-056908.
PR 22-AUG-1997; US-056909.
PR 22-AUG-1997; US-056910.
PR 22-AUG-1997; US-056911.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057659.
PR 05-SEP-1997; US-057661.
PR 12-SEP-1997; US-058785.
PA (HUMA-) HUMAN GENOME SCI INC.

PI Bednarik DP, Brewer LA, Carter KC, Duan R, Eder R, Endress GA,
PI Feng P, Ferlie AM, Fischer CL, Florence KA, Greene JM, Hu JS,
PI Kyaw H, Lafleur DM, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
DR WPI: 98-506364/43.
DR P-PSDB: W74786.
PT New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1; Page 289-290; 721pp; English.
CC This sequence represents a nucleic acid molecule designated Gene 56 from
CC the human cDNA clone HSAS65 (deposited as clone ATCC 97899 and ATCC
CC 209045) which encodes a secreted human protein. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC re portion (e.g. V59502) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see V59511 for described
CC uses).
SQ Sequence 732 BP; 220 A; 167 C; 183 G; 162 T;

Query Match 1.2%; Score 37; DB 1; Length 732;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 674 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 710

RESULT 12
ID N50414 standard; DNA; 970 BP.
AC N50414;
DT 15-JUN-1992 (first entry)
DE Sequence encoding pig elastase I.
KW Arteriosclerosis; E.coli; ds.
OS Sus scrofa domestica.
FH Key Location/Qualifiers
FT cds 1..801
FT /*tag- a
PN J60186284-A.
PD 21-SEP-1985.
PF 07-MAR-1984; 043155.
PR 07-MAR-1984; JP-043155.
PA (KIRI) KIRIN BREWERY KK.
DR WPI: 85-273645/44.
DR P-PSDB: P50360.
PT Biochemical prepn. of pig-elastase I - uses DNA chain to produce
PT plasmid to transform Escherichia coli or Saccharomyces
PT microorganism.
PS Claim 1; Fig 1; 16pp; Japanese.
CC The gene product may be expressed from an E.coli or Saccharomyces
CC host transformed with a plasmid encoding the pig elastase I gene.
CC Elastase I is useful in the treatment of arteriosclerosis, and may be
CC efficiently produced by this technique.
SQ Sequence 970 BP; 247 A; 261 C; 262 G; 200 T;

Query Match 1.2%; Score 37; DB 1; Length 970;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 932 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 968

RESULT 13
ID T73851 standard; cDNA; 1001 BP.
AC T73851;
DT 08-SEP-1997 (first entry)
DE Full length 3' UTR of prohibitin wild-type B allele.
KW Prohibitin; tumor suppressor gene; loss of heterozygosity; LOH;
KW cell immortalisation; complementation group B; mutation; 3' UTR;
KW untranslated region; diagnosis; susceptibility; cancer; breast;
KW screen; genotype; treatment; early stage tumour; ds.
OS Synthetic.
PN W09640919-A1.
PD 19-DEC-1996.
PF 07-JUN-1996; U09687.
PR 07-JUN-1995; US-473486.
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
PI Dellorco RT, Jupe ER, King RL, Liu X, McClung JK;
DR WPI: 97-052321/05.
PT Prohibitin gene 3' untranslated region - reintroduction of which
PT into early stage tumours is useful in cancer treatment
PS Example 1; Page 41-42; 60pp; English.

CC Proliferative activity, is postulated to be a tumour suppressor gene
CC whose expression, when lost, contributes to the immortalisation of cells
CC from one or more of the four complementation groups proposed by Pereira
CC and is apparently involved in the process of immortalisation in a
CC group of human cells that are classified in complementation group B.
CC Mutations in the 3' untranslated region (3' UTR) of the B type allele are
CC diagnostic for increased susceptibility to cancer, particularly breast
CC cancer. Diagnosis comprises determining an individual's prohibitin
CC genotype and then screening for alterations in the 3' UTR of the
CC prohibitin gene. Also cancer can be treated by reintroducing at least a
CC portion of the wild type 3' UTR of a prohibitin gene into early stage
CC tumours. The present sequence is that of the full 3' UTR. The cDNA and
CC genomic sequences are identical, showing that the full 3' UTR is
CC transcribed from the genome without introns.
SQ Sequence 1001 BP; 267 A; 253 C; 250 G; 231 T;

Query Match 1.2%; Score 37; DB 1; Length 1001;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 940 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 976

RESULT 14
ID V69433 standard; cDNA; 1889 BP.
AC V69433;
DT 15-MAR-1999 (first entry)
DE Human edg-6 cDNA #2.
KW Endothelial differentiation gene receptor homologue; human; HRDG;
KW EDG-6; detection; diagnosis; drug screening; treatment; inflammation;
KW disease; viral; bacterial; fungal infection; allergic response; injury;
KW hereditary disease; lymphoma; carcinoma; lymphoid; neuronal; cascade;
KW lymphocyte trafficking; leukocyte trafficking; signal transduction;
KW rheumatoid synovium; autoimmune disorder; ss.
OS Homo sapiens.
PN W09835062-A1.
PD 26-NOV-1998.
PF 22-MAY-1998; CA0487.
PR 22-MAY-1997; US-861747.
PA (ALIX) ALTELIX BIOPHARMACEUTICALS INC.
PI Munroe DG, Vyas TB;
DR WPI: 99-070147/06.
PT New isolated human endothelial differentiation gene-6 receptor
PT homologue - used to develop products for treating e.g. infections,

PT allergic responses, trauma, hereditary diseases, lymphoma, carcinoma
 PT or autoimmune disorders
 PS Disclosure: Fig 2; 59pp; English.
 CC This sequence encodes a novel human endothelial differentiation gene
 CC (EDG)-6 receptor homologue designated HEDG. The encoded protein can be
 CC used for detection, diagnosis and drug screening. The protein can also
 CC be used to treat inflammation or diseases including viral, bacterial, or
 CC fungal infections, allergic responses, mechanical injury associated with
 CC trauma, hereditary diseases, lymphoma or carcinoma, or other conditions
 CC which activate the genes of lymphoid or neuronal tissues. It can also be
 CC used for treating problems involving excessive lymphocyte and leukocyte
 CC trafficking. Inhibitors of HEDG are useful for controlling signal
 CC transduction and signalling cascades in cells of the rheumatoid synovium.
 CC The protein can also be used for treating autoimmune disorders such as
 CC myasthenia gravis.
 SQ Sequence 1889 BP; 360 A; 575 C; 537 G; 416 T;

Query Match 1.2%; Score 37; DB 1; Length 1889;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||||||||||||||||||||||||||||||
 Db 1845 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 1881

RESULT 15
 Q25156
 ID Q25156 standard; cDNA; 2158 BP.
 AC Q25156;
 DT 18-NOV-1992 (first entry)
 DE Alpha-GalNAC from pAGB-3.
 KM Lysosome; Schindler disease; infantile neuroaxonal dystrophy; ss.
 OS Homo sapiens.
 FH Key
 FT cds
 FT 345..1580
 FT /tag= a
 FT /label= alpha-GalNAC
 FT 345..395
 FT /tag= b
 FT mat_peptide 396..1580
 FT /tag= c
 FT 2073..2078
 FT /tag= d
 FT polya_signal 2025..2029
 FT protein_bind
 FT /tag= e
 FT /note= "recognised by the U4 small nuclear
 FT ribonucleoprotein"
 FT
 PN W09207936-A.
 PD 14-MAY-1992.
 PR 23-OCT-1991; U07872.
 PR 24-OCT-1990; US-602608.
 PA (MOUN) MOUNT SINAI SCHOOL MEDICINE.
 PI Bishop DF, Desnick RJ, Ioannou YA, Wang AM;
 DR WPI: 92-183672/22.
 DR P-PSDB: R24291.
 PT Cloning and expression of alpha-n-acetyl-galactose aminidase -
 PT used in enzyme replacement therapy for Schindler disease
 PS Disclosure: Fig 2 (A-D); 71pp; English.
 CC The sequence is of the pAGB-3 cDNA insert contg. the complete coding
 CC region for human alpha-GalNAC.
 CC The availability of the full length cDNA for alpha-GalNAC allows
 CC the study of the genomic organisation and evolution of this
 CC lysosomal gene, and the characterisation of molecular lesions
 CC causing Schindler disease.
 SQ Sequence 2158 BP; 517 A; 610 C; 576 G; 455 T;

Query Match 1.2%; Score 37; DB 1; Length 2158;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||||||||||||||||||||||||||||||
 Db 2090 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 2126

RESULT 16
 ID V55872
 ID V55872 standard; cDNA; 2702 BP.
 AC V55872;
 DT 18-NOV-1998 (first entry)
 DE Plant acetoacetate synthase (ALS) large subunit protein encoding cDNA.
 KM ALS; small subunit; acetoacetate synthase; plant; ssuALS; lsuALS;
 KM Nicotiana glumabaginfolia; plasmid vector; herbicide; holoenzyme;
 KM large subunit; ss.
 OS Nicotiana glumabaginfolia.
 FH Key
 FT CDS
 FT 492..2492
 FT /tag= a
 FT /product= "ALS large subunit protein"

PN W09837206-A1.
 PD 27-AUG-1998.
 PR 23-FEB-1998; U03506.
 PR 24-FEB-1997; US-039148.
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 PI Abell LM, Hershey HP;
 DR WPI: 98-467568/40.
 DR P-PSDB: W79141.
 PT New nucleic acid encoding a plant aceto: lactate synthase small
 PT sub:unit - that combines with the large sub:unit to give a
 PT holoenzyme having higher activity than the large sub:unit alone
 PS Example 6; Pages 36-39; 47pp; English.
 CC This cDNA encodes the large subunit of a plant aceto: lactate synthase.
 CC The invention provides a cDNA sequence contained in the plasmid vector
 CC encoding the small subunit of ALS (ssuALS). The plasmid vector comprising
 CC the ssuALS nucleotide operably linked to a regulatory sequence can be
 CC used to transform host cells for the recombinant production of the ssuALS
 CC protein which is used for evaluating a compound for aceto: lactate synthase
 CC inhibition and so for selecting potential herbicides. Evaluation of a
 CC compound for aceto: lactate synthase (ALS) inhibition comprises expression
 CC and purification of plant ssuALS from the transformed host, mixing this
 CC ssuALS with the large subunit of ALS to form a holoenzyme which is then
 CC treated with a test compound. Treated and untreated holoenzyme activity
 CC are compared to select compounds with potential for herbicidal activity.
 CC Previously, large subunits of plant ALS (lsuALS) have been isolated,
 CC but the existence of ssuALS had not been verified. Mixing the two
 CC subunits results in a holoenzyme that has a 4-15 fold increase in
 CC specific activity over lsuALS alone.
 SQ Sequence 2702 BP; 668 A; 645 C; 642 G; 747 T;

Query Match 1.2%; Score 37; DB 1; Length 2702;
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3116
 |||||||||||||||||||||||||||||||||
 Db 2660 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 2696

RESULT 17
 ID X51737
 ID X51737 standard; DNA; 340 BP.
 AC X51737;
 DT 17-JUN-1999 (first entry)
 DE DNA encoding a human secreted protein.
 KM Human secreted protein; cancer; immune disorder; infection;
 KM inflammatory disorder; skin disorder; tumour; atherosclerosis;
 KM restenosis; autoimmune disorder; Alzheimer's disease;
 KM peripheral neuropathy; trauma; spinal cord injury; allergy;
 KM hemopoietic disorder; skeletal disorder; neurological disorder;
 KM arthritic disorder; asthma; immunodeficiency disease; AIDS;
 KM transplant rejection; ss.
 OS Homo sapiens.

PN W09911293-A1.
 PD 11-MAR-1999.
 PF 03-SEP-1998; U18360.
 PR 12-SEP-1997; US-058974.
 PR 05-SEP-1997; US-057626.
 PR 05-SEP-1997; US-057663.
 PR 05-SEP-1997; US-057669.
 PR 12-SEP-1997; US-058666.
 PR 12-SEP-1997; US-058667.
 PR 12-SEP-1997; US-058973.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Brewer LA, Edner R, Lafleur DM, Moore PA, Olsen HS,
 PI Rosen GA, Ruben SM, Shi Y;
 DR WPI; 99-20498/17.
 DR P-PSDB; Y12950.
 PT New isolated human genes and the secreted polypeptides they encode
 PT - useful for diagnosis and treatment of e.g. neurological disorders,
 PT tumours, immune disorders, inflammation or haematological disorders
 PS Claim 1; Page 172; 215pp; English.
 CC X51701-55 encode human secreted proteins. The polynucleotides and
 CC their corresponding secreted polypeptides are useful for preventing,
 CC treating or ameliorating medical conditions, e.g. by protein or gene
 CC therapy. Pathological conditions can also be diagnosed by determining
 CC the amount of the new polypeptides in a sample or by determining the
 CC presence of mutations in the new polynucleotides. Specific uses are
 CC described for each polynucleotide, based on which tissues they are
 CC most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, immune disorders, infection,
 CC inflammatory disorders, skin disorders, tumours, atherosclerosis,
 CC restenosis, autoimmune disorders, Alzheimer's disease, peripheral
 CC neuropathies, trauma, spinal cord injuries, allergy, hematopoietic
 CC disorders, skeletal disorders, neurological disorders, arthritic
 CC disorders, asthma, immunodeficiency diseases, AIDS and transplant
 CC rejection. The polypeptides are also useful for identifying their
 CC binding partners.
 CC Sequence 340 BP; 133 A; 69 C; 73 G; 65 T;
 SQ

Query Match 1.2%; Score 36; DB 1; Length 340;
 Best Local Similarity 100.0%; Pred. No. 7.6e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 ||||||||||||||||||||||||||||||||
 DB 267 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 302

RESULT 18
 V90045/C
 ID V90045 standard; cDNA: 378 BP.
 AC V90045;
 DT 15-FEB-1999 (first entry)
 DE EST clone CM1550.
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 OS Homo sapiens.
 PN W09845436-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998; U06955.
 PR 10-APR-1997; US-838821.
 PA (GEMV) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 PI Reicle LA, Spaulding V, Treacy M;
 DR WPI; 99-070077/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.
 PS Claim 1; Page 414; 618pp; English.
 CC The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make

CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 CC Sequence 378 BP; 76 A; 99 C; 92 G; 107 T;
 SQ

Query Match 1.2%; Score 36; DB 1; Length 378;
 Best Local Similarity 100.0%; Pred. No. 7.5e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 ||||||||||||||||||||||||||||||||
 DB 323 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 288

RESULT 19
 ID Q23343 standard; cDNA: 433 BP.
 AC Q23343;
 DT 12-AUG-1992 (first entry)
 DE Murine CP-10.
 KW Chemotactic protein; neutrophil; monocyte; macrophage;
 KW Inflammation; autoimmune disease; delayed hypersensitivity;
 KW CF; cystic fibrosis; emphysema; diagnosis; ss.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT cds 52..321
 FT /*tag= a
 FT /product= CP-10
 FT
 PN W09204376-A.
 PD 19-MAR-1992.
 PF 05-SEP-1991; AU0410.
 PR 05-SEP-1990; AU-002127.
 PR 05-FEB-1991; AU-004463.
 PA (HEAR-) HEART RES INST LTD.
 PI Geerzy C, Simpson RJ, Lackmann M;
 DR WPI; 92-114301/14.
 DR P-PSDB; R22429.
 PT New chemotactic protein CP-10 for altering inflammatory capacity
 PT - controls inflammatory conditions, cystic fibrosis and
 PT emphysema, antagonists for detecting CP-10
 PS Claim 18; Page 38; 56pp; English.
 CC The mCP10 cDNA was derived from a Sepharose-Concanavalin A-activated
 CC spleen cell library prep. using cells from A/J mice. CP-10 has an
 CC apparent mol. wt. of ca. 10 kD and is chemotactic for neutrophils,
 CC monocytes/macrophages and/or other mammalian cells. CP-10 is useful
 CC (e.g. when admin. locally around a tumour or site of infection) to
 CC attract neutrophils etc., esp. in immunocompromised patients or
 CC patients allergic to skin test antigens. Antibodies to CP-10 are
 CC useful for control of inflammation and/or tissue damage associated
 CC with inflammation, (e.g. due to autoimmune disease or delayed
 CC hypersensitivity), cystic fibrosis or emphysema. Ab can also be used
 CC to purify CP-10 and in diagnosis.
 CC Sequence 433 BP; 161 A; 83 C; 87 G; 102 T;
 SQ

Query Match 1.2%; Score 36; DB 1; Length 433;
 Best Local Similarity 100.0%; Pred. No. 7.3e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 ||||||||||||||||||||||||||||||||
 DB 387 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 422

RESULT 20
 X22245

ID X22245 standard; DNA; 437 BP.
AC X22245;
DE 18-MAY-1999 (first entry)
KW Human secreted protein gene 35 clone HMADU73.
KW Human; secreted protein; gene therapy; protein therapy; cancer; weight;
KW tumour; chromosome mapping; forensic; haematological disease; allergy;
KW inflammation; cell proliferation; viral infection; wound healing;
KW modulation; appetite; behaviour; food additive; preservative; ss.
OS Homo sapiens.
PN WO903990-A1.
PD 28-JAN-1999.
PE 15-JUL-1998: U14613.
PR 18-AUG-1997: US-056361.
PR 16-JUL-1997: US-052661.
PR 16-JUL-1997: US-052870.
PR 16-JUL-1997: US-052871.
PR 16-JUL-1997: US-052872.
PR 16-JUL-1997: US-052873.
PR 16-JUL-1997: US-052874.
PR 16-JUL-1997: US-052875.
PR 22-JUL-1997: US-053440.
PR 22-JUL-1997: US-053441.
PR 22-JUL-1997: US-053442.
PR 18-AUG-1997: US-055683.
PR 18-AUG-1997: US-055724.
PR 18-AUG-1997: US-055725.
PR 18-AUG-1997: US-055726.
PR 18-AUG-1997: US-055946.
PR 18-AUG-1997: US-055952.
PR 18-AUG-1997: US-055985.
PR 18-AUG-1997: US-055989.
PR 18-AUG-1997: US-056359.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Duan R, Feng P, Ferlie AM, Florence KA, Fouad J,
PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE,
PI Yu G;
DR WPI: 99-132234/11.
DR P-PSDB: Y01417.
PT New nucleic acids encoding secreted human proteins - potentially
PT useful for treating and diagnosing diseases and identifying specific
PT binding agents
PS Claim 4: Page 187: 251pp: English.
CC The invention relates to nucleic acid sequences (X22211 to X22283)
CC encoding human secreted proteins (Y01383 to Y01454). The secreted protein
CC gene sequences are deposited with the ATCC under deposit number ATCC
CC 209138, 209139 or 209141. Host cells containing vectors comprising the
CC nucleic acid sequences are used for the recombinant expression of the
CC secreted proteins. The polynucleotide and amino acid sequences are useful
CC for preventing, creating or ameliorating medical conditions e.g. by
CC protein or gene therapy. Pathological conditions can be also diagnosed by
CC determining the amount of the new polypeptides in a sample or by the
CC presence of mutations in the new polynucleotides. The nucleic acid
CC sequences, or its fragments, are useful for chromosome identification and
CC mapping; as antisense and triplex-forming therapeutics; in gene therapy;
CC for (forensic) identification of individuals; as molecular weight
CC markers; to identify related sequences or specific mRNA; in preparation
CC of oligomers and to raise anti-DNA antibodies. Antibodies are useful as
CC immunosay reagents (including for in vivo imaging) and therapeutically
CC to inhibit or activate particular polypeptides. A very wide range of
CC disorders may be treated with the polynucleotide and polypeptide
CC sequences, e.g. autoimmune or haematological diseases, allergy,
CC inflammation, cancer or other forms of cell proliferation, viral or other
CC infections. The sequences may also be useful in wound healing, to
CC modulate differentiation of embryonic stem cells, to modulate weight,
CC appetite, behaviour etc. and as food additive or preservative. The
CC present sequence represents a gene encoding a human secreted protein
CC (see descriptor line for gene number and clone identification).
SQ Sequence 437 BP; 136 A; 128 C; 97 G; 72 T;

Query Match 1.2%; Score 36; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 380 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415
RESULT 21
ID X04408 standard; DNA; 505 BP.
AC X04408;
DE 13-APR-1999 (first entry)
KW Human secreted protein gene.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9856804-A1.
PD 17-DEC-1998.
PE 11-JUN-1998: U12125.
PR 02-OCT-1997: US-061060.
PR 13-JUN-1997: US-049547.
PR 13-JUN-1997: US-049548.
PR 13-JUN-1997: US-049549.
PR 13-JUN-1997: US-049550.
PR 13-JUN-1997: US-049606.
PR 13-JUN-1997: US-049607.
PR 13-JUN-1997: US-049608.
PR 13-JUN-1997: US-049609.
PR 13-JUN-1997: US-049610.
PR 13-JUN-1997: US-049611.
PR 13-JUN-1997: US-050566.
PR 13-JUN-1997: US-050901.
PR 13-JUN-1997: US-052989.
PR 08-JUL-1997: US-051919.
PR 18-AUG-1997: US-055984.
PR 12-SEP-1997: US-058665.
PR 12-SEP-1997: US-058668.
PR 12-SEP-1997: US-058669.
PR 12-SEP-1997: US-058750.
PR 12-SEP-1997: US-058971.
PR 12-SEP-1997: US-058972.
PR 12-SEP-1997: US-058975.
PR 02-OCT-1997: US-060834.
PR 02-OCT-1997: US-060841.
PR 02-OCT-1997: US-060844.
PR 02-OCT-1997: US-060865.
PR 02-OCT-1997: US-061059.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Ferlie AM, Feng P, Greene JM, Lafleur DM,
PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P,
PI Yu G;
DR WPI: 99-080881/07.
DR P-PSDB: W78223.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1: Page 262: 380pp: English.
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number, and the clone it is derived from, are
CC detailed in the descriptor line. The gene can be used to generate fusion
CC proteins by linking to the gene to a human immunoglobulin Fc portion
CC (e.g. X04302) for increasing the stability of the fused protein as
CC compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic acid
CC sequences: X04311-X04410, amino acid sequences W78126-W78225) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample

CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on
CC which tissues they are most highly expressed in (see X04311 for described
CC uses).

SO Sequence 505 BP; 135 A; 150 C; 120 G; 97 T;

Query Match 1.2%; Score 36; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 7; le-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 458 TCAAAAAAAAAAAAAAAAAAAAAA 493

RESULT 22

T64563 ID T64563 standard; cDNA to mRNA; 558 BP.

AC T64563;

DT 01-DEC-1997 (first entry)

DE Mouse thymus and activation regulated cytokine (TARC) cDNA.

KM Mouse; thymus; activation; regulation; chemokine; TARC; CC type;

KM peripheral; blood; monocyte; CCR4; T cell; lymphocyte;

OS Immunomodulation; inflammation; murine; ds.

PI Mus spp.

FT Key Location/Qualifiers

FT cds 2..283

FT /*tag= a

PN MO9711969-A1.

PD 03-APR-1997.

PE 27-SEP-1996; J02801.

PR 13-MAR-1996; JP-056044.

PR 27-SEP-1995; JP-249457.

PA (SHIO) SHIONOGI & CO LTD.

PI Imai T, Yoshida T, Yoshie O;

DR WPI; 97-212853/19.

DR P-PSDB; W14918.

PT New CC-chemokine-like protein - is expressed by peripheral blood

PT monocytes under immunological stimulus, useful as immunomodulator

PT and antiinflammatory agent

PS Claim 12; Pages 59-60; 102pp; Japanese.

CC The present sequence encodes mouse thymus and activation regulated

CC chemokine (TARC), which is similar to CC type chemokines and

CC expressed in peripheral blood monocytes under immunological

CC stimulus, e.g. phyto-haemagglutinin or soluble cytokine, and the

CC thymus, but not by the spleen. TARC is active against CCR4

CC expressing cells, e.g. T cells, peripheral lymphocytes and

CC activated peripheral T cells, especially Hut 78, Hut 102, Jurkat,

CC MT2 and MT4 cells, and is useful as an immunomodulator or

CC anti-inflammatory agent.

SO Sequence 558 BP; 210 A; 124 C; 121 G; 103 T;

Query Match 1.2%; Score 36; DB 1; Length 558;

Best Local Similarity 100.0%; Pred. No. 7e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 464 TCAAAAAAAAAAAAAAAAAAAAAA 499

RESULT 23

T69172 ID T69172 standard; cDNA; 581 BP.

AC T69172;

DT 12-JAN-1998 (first entry)

DE Trypanosoma cruzi antigen (ubiquitin) cDNA.

KM Antigen; epitope; vaccine; protective immunity; Chagas disease;

KM diagnosis; therapy; immunoassay; ubiquitin; ss.

OS Trypanosoma cruzi Tulane strain C2.

FT Key Location/Qualifiers

FT CDS 45..431

FT /*tag= a

PN MO9718475-A1.

PD 22-NOV-1997.

PE 14-NOV-1996; U18624.

PR 14-NOV-1995; US-557309.

PA (CORI-) CORIXA CORP.

PI Houghton RL, Lodes MJ, Reed SG, Skeiky YAM;

DR WPI; 97-289413/26.

DR P-PSDB; W26534.

PT Diagnosing Trypanosoma cruzi infection by detecting antibodies to

PT novel antigens - which are useful in vaccines to provide protective

PT immunity against Chagas' disease

PS Claim 1; Page 50-51; 110pp; English.

CC This novel DNA molecule was isolated by screening a Trypanosoma

CC cruzi cDNA expression library with pools of sera from infected

CC individuals. It encodes T. cruzi ubiquitin (see W26534). 22

CC Isolated genomic DNA of cDNA molecules (T69151-72) encode full-

CC length antigens (see W26530-41), or epitope-containing repeat

CC sequences (see W19094-102, W19079-86 and W26542-44) of native

CC antigens, that can be used in a variety of immunoassays for

CC detecting T. cruzi infection in a blood, serum, plasma, saliva,

CC cerebrospinal fluid or urine sample. The polypeptides are also

CC useful in vaccines and pharmaceutical compositions for inducing

CC protective immunity against Chagas disease. The isolated DNA

CC molecules can be used for recombinant production of the antigenic

CC polypeptides.

SO Sequence 581 BP; 160 A; 139 C; 155 G; 127 T;

Query Match 1.2%; Score 36; DB 1; Length 581;

Best Local Similarity 100.0%; Pred. No. 7e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 543 TCAAAAAAAAAAAAAAAAAAAAAA 578

RESULT 24

T64562 ID T64562 standard; cDNA to mRNA; 582 BP.

AC T64562;

DT 01-DEC-1997 (first entry)

DE Human thymus and activation regulated cytokine (TARC) cDNA.

KM Human; thymus; activation; regulation; chemokine; TARC; CC type;

KM peripheral; blood; monocyte; CCR4; T cell; lymphocyte;

OS Immunomodulation; inflammation; ds.

PI Homo sapiens.

FT Key Location/Qualifiers

FT cds 53..337

FT /*tag= a

PN MO9711969-A1.

PD 03-APR-1997.

PE 27-SEP-1996; J02801.

PR 13-MAR-1996; JP-056044.

PR 27-SEP-1995; JP-249457.

PA (SHIO) SHIONOGI & CO LTD.

PI Imai T, Yoshida T, Yoshie O;

DR WPI; 97-212853/19.

DR P-PSDB; W14917.

PT New CC-chemokine-like protein - is expressed by peripheral blood

PT monocytes under immunological stimulus, useful as immunomodulator

PT and antiinflammatory agent

PS Claim 10; Pages 58-59; 102pp; Japanese.

CC The present sequence encodes human thymus and activation regulated

CC chemokine (TARC), which is similar to CC type chemokines and

CC expressed in peripheral blood monocytes under immunological

CC stimulus, e.g. phyto-haemagglutinin or soluble cytokine, and the

CC thymus, but not by the spleen. TARC is active against CCR4

CC expressing cells, e.g. T cells, peripheral lymphocytes and

CC activated peripheral T cells, especially Hut 78, Hut 102, Jurkat,

CC MT2 and MT4 cells, and is useful as an immunomodulator or

CC anti-inflammatory agent.
SQ Sequence 582 BP; 162 A; 168 C; 149 G; 103 T;
Query Match 1.2%; Score 36; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 537 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 572
RESULT 25
X20419
ID X20419 standard; DNA; 602 BP.
AC X20419;
DT 04-MAY-1999 (first entry)
DE Human secreted protein, gene 8.
KW Human secreted protein; fusion protein; gene therapy; protein therapy;
diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
inflammation; ischaemic shock; Alzheimer's disease; osteoporosis; AIDS;
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
KM Homo sapiens.
OS Homo sapiens.
PN MO9906423-A1.
PD 11-FEB-1999.
PF 29-JUL-1998; U15949.
PR 19-AUG-1997; US-056730.
PR 30-JUL-1997; US-054209.
PR 30-JUL-1997; US-054211.
PR 30-JUL-1997; US-054212.
PR 30-JUL-1997; US-054213.
PR 30-JUL-1997; US-054214.
PR 30-JUL-1997; US-054215.
PR 30-JUL-1997; US-054217.
PR 30-JUL-1997; US-054218.
PR 30-JUL-1997; US-054234.
PR 30-JUL-1997; US-054236.
PR 18-AUG-1997; US-055968.
PR 18-AUG-1997; US-055969.
PR 18-AUG-1997; US-055972.
PR 19-AUG-1997; US-056534.
PR 19-AUG-1997; US-056534.
PR 19-AUG-1997; US-056554.
PR 19-AUG-1997; US-056561.
PR 19-AUG-1997; US-056727.
PR 19-AUG-1997; US-056729.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Carter KC, Endress GA, Fan P, Feng P, Kyaw H, Lafleur DM,
PI Li Y, Moore PA, Rosen CA, Ruben SM, Shi Y, Wei Y,
PI Zeng Z,
PI WPI; 99-153691/13.
DR P-PSDB; Y00265.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1, Page 214, 312pp; English.
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number is given in the descriptor line. The gene
CC can be used to generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. X20403) for increasing the stability of
CC the fused protein as compared to the human protein only.
CC The invention relates to 83 novel genes and their fragments (nucleic acid
CC sequences: X20412-X20499; amino acid sequences Y00258-Y00377) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on

CC which tissues they are most highly expressed in (see X20412 for described
CC uses).
SQ Sequence 602 BP; 194 A; 133 C; 152 G; 123 T;
Query Match 1.2%; Score 36; DB 1; Length 602;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 528 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 563
RESULT 26
Q72956
ID Q72956 standard; CDNA; 684 BP.
AC Q72956;
DT 04-JUN-1995 (first entry)
DE Clone pSP65-Xai-11 encoding recombinant Factor Xa inhibitor (FXai).
KW Leech, factor Xa inhibitor; ss.
OS Hirudo medicinalis.
FH Key Location/Qualifiers
FT cds 139..684
FT misc_difference 1..684
FT /*tag- a
FT /*tag- b
FT /*label= N= unspecified base
PN MO9423709-A.
PD 27-OCT-1994.
PF 08-APR-1994; U03918.
PR 09-APR-1993; US-045804.
PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
PA (YISS) YISSUM RES & DEV CO.
PI Fischer M, Goldlust A, Guy R, Levanon A, Panet A,
PI Rigbi M, Werber MM, Zeelon EP;
DR WPI; 94-341657/42.
DR P-PSDB; R62619.
PT Recombinant factor 10a inhibitor of Hirudo medicinalis - for
PT treating excessive blood coagulation, partic. thrombosis, also
PT related DNA, vectors, transformed cells and antibodies
PS Claim 3, Figure 10; 107pp; English.
CC Total RNA was extracted from 120 leeches and used to prepare a
CC cDNA library which was screened for clones contg. FXai DNA using
CC as probe radiolabeled DNA from the PCR derived plasmid pSP65-Xai-
CC 4 (see Q72957). The plasmid of one of the positive clones - pSK-
CC Xai-11 (clone 11) contains a 684 bp insert (see Q72956) essentially
CC identical to the previously obt. PCR derived sequences but contg.
CC 78 additional bps at its 5' end. It encodes a prepeptide of 110
CC AAs (incl. the N-terminal Met) extending from Met 1 to Gly 110 (see
CC R62619). AAs 1-25 constitute a leader sequence, and the mature
CC protein comprises Tyr 26 to Gly 110. The sequence Tyr 2 - Gly 86
CC of clone 4 (Q72957) is equiv. to the sequence Tyr 26 - Gly 110
CC of clone 11 (Q72956). There are only two differences between
CC clones 4 and 11. The polypeptide expressed by clone 11 is
CC referred to as recombinant FXai. A plasmid comprising the DNA
CC designated pSP65-Xai-11 and deposited under ATCC Accession No.
CC 69138 is claimed.
SQ Sequence 684 BP; 263 A; 110 C; 120 G; 190 T;
Query Match 1.2%; Score 36; DB 1; Length 684;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 636 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 671
RESULT 27
X30351
ID X30351 standard; DNA; 687 BP.

AC X30351;
DT 14-MAY-1999 (first entry)
DE DNA encoding a human secreted protein.
KW Secreted protein; cancer; neurodegenerative disorder;
KW developmental abnormality; foetal deficiency; blood disorder;
KW CNS disorder; immune system disease; autoimmune disease; hepatic disease;
KW renal disease; diabetes; inflammation; allergy; ischemic shock;
KW Alzheimer's; cognitive disorder; schizophrenia; cardiovascular disorder;
KW prostate disease; asthma; osteoporosis; arthritis; ss.
OS Homo sapiens.
PN M09907891-A1.
PD 18-FEB-1999.
PF 04-AUG-1998; U16235.
PR 19-AUG-1997; US-036732.
PR 05-AUG-1997; US-054798.
PR 05-AUG-1997; US-054803.
PR 05-AUG-1997; US-054804.
PR 05-AUG-1997; US-054806.
PR 05-AUG-1997; US-054807.
PR 05-AUG-1997; US-054808.
PR 05-AUG-1997; US-054809.
PR 05-AUG-1997; US-055309.
PR 05-AUG-1997; US-055310.
PR 05-AUG-1997; US-055312.
PR 05-AUG-1997; US-055386.
PR 05-AUG-1997; US-055311.
PR 18-AUG-1997; US-055970.
PR 18-AUG-1997; US-055985.
PR 19-AUG-1997; US-055365.
PR 19-AUG-1997; US-056366.
PR 19-AUG-1997; US-056557.
PR 19-AUG-1997; US-056370.
PR 19-AUG-1997; US-056371.
PR 19-AUG-1997; US-056563.
PR 19-AUG-1997; US-056731.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Ferrie AM, Greene JM, Janat F, Ni J,
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR, Young PE, Yu G;
PI WPI: 99-167452/14.
DR P-PSDB: Y10830.
PT New isolated human genes encoding secreted polypeptides - useful for
PT diagnosis and treatment of pathological diseases
PS Claim 3; Page 247; 331pp; English.
CC The specification describes secreted proteins and their corresponding
CC polynucleotides which are useful for preventing, treating or ameliorating
CC medical conditions, e.g. by protein or gene therapy. Pathological
CC conditions can also be diagnosed by determining the amount of the
CC secreted polypeptides in a sample or by determining the presence of
CC mutations in the polynucleotides. Specific uses are described for each
CC of the products, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumours, neurodegenerative disorders, developmental
CC abnormalities and foetal deficiencies, blood disorders, CNS disorders,
CC diseases of the immune system, autoimmune diseases, hepatic and renal
CC disease, diabetes, inflammation, allergies, ischemic shock, Alzheimer's
CC and cognitive disorders, schizophrenia, cardiovascular disorders,
CC prostate diseases, asthma, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, diseases of testes, lung or
CC thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
SQ Sequence 687 BP; 236 A; 123 C; 152 G; 173 T;

Query Match 1.2%; Score 36; DB 1; Length 687;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 634 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 669

RESULT 28

Q22005
ID Q22005 standard; DNA; 742 BP.
AC Q22005;
DT 09-JUN-1992 (first entry)
DE Sequence encoding Cor a I allergen of hazel.
KW Fagales; hazel; IgE; tree allergen; pollen; antibodies; T-cell
KW response; ss.
OS Fagales hazel.
FH Key cds Location/Qualifiers
FT 1..483
FT /*tag= a
FT /*product= Aln_g-I
FT 692..742
FT /*tag= b
polya_site
FT
FT W09202621-A.
PN 06-FEB-1992.
PD 06-MAY-1991; E01479.
PF 08-AUG-1990; AT-001688.
PR 11-APR-1991; US-683831.
PR (BIOM-) BIOMAY BIOECHAN PRO.
PI Breitenbach H, Reikertorfer A, Valenta R, Hoffmann-Sommergruber K;
PI Breitenbach H, Kraft D, Rumpold H, Scheiner O;
DR WPI: 92-080075/10.
DR P-PSDB: R21793.
DT Aln g I, Cor a I and Bet v I allergens - and DNA from alder,
PT hazel and birch, useful in diagnosis or therapy of allergic
PT diseases
PS Claim 8; Page 28; 54pp; English.
CC The cDNA sequence encoding the Cor a I allergen of hazel was obt
CC from the polyA+ mRNA isolated from ripe hazel pollen. The resulting
CC cDNA was amplified by PCR using primers whose sequences were derived
CC from the N-terminal amino-acid sequence of Cor a I. The DNA fragment
CC produced was cloned into the pBluescript KS vector and transformed
CC into E. coli XL1-Blue cells. The sequence allows mammals to be
CC tested for allergic reactions to specific tree allergens. The
CC derived polypeptide may be used to challenge the mammal to elicit
CC bronchial, conjunctival, dermal, nasal or oral provocation. The
CC polypeptide may be used to treat a mammal afflicted with a pollen
CC allergy. It is administered in an amount sufficient to hyposenstise
CC the mammal to Cor a I.
CC See also Q22001-8 and Q21974.
SQ Sequence 742 BP; 250 A; 134 C; 172 G; 186 T;

Query Match 1.2%; Score 36; DB 1; Length 742;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 691 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 726

RESULT 29
ID T59912 standard; cDNA; 780 BP.
AC T59912;
DT 04-JUN-1997 (first entry)
DE Human transcription regulatory factor SRB7 cDNA.
KW Transcription regulatory factor; suppressor of RNA polymerase B;
KW SRB7; RNA polymerase II; holoenzyme; SWI/SNF; ss.
OS Homo sapiens.
FH Key cds Location/Qualifiers
FT 10..444
FT /*tag= a
FT 627
FT allele
FT /*tag= b
FT /note= "there is a polymorphism at position 627;
FT in some individuals it is GATC (a
FT restriction site for Sau3A), in others it
FT is GATT"
PN W09708301-A1.
PD 06-MAR-1997.

PF 28-AUG-1996; U14192.
PR 31-AUG-1995; US-521872.
PR 11-OCT-1995; US-540804.
PR 26-JAN-1996; US-590399.
PA (WHEB) WHITEHEAD INST BIOMEDICAL RES.
PI Chao DM, Koleske AJ, Thompson CM, Young RA;
DR WPI: 97-179258/16.
DR P-PSDB: W13829.
PT Purified RNA polymerase II holo:enzyme - comprises RNA polymerase II
PT and one or more regulatory proteins, pref. suppressor of RNA
PT polymerase B proteins or SWI/SNF proteins
PS Claim 11; Fig 17a; 154pp; English.
CC A cDNA clone (T59912) codes for the human homologue (W13829),
CC hSRB7, of the yeast transcription regulatory factor ySRB7 (see
CC also W13824). It was cloned and sequenced using 3 overlapping
CC expressed sequence tags. hSRB7 is 35% identical to ySRB7,
CC complements a ySRB7 deletion and, like its yeast counterpart, binds
CC to the C-terminal domain of RNA polymerase II. It forms part of a
CC holoenzyme complex that supports active transcription. SRB
CC nucleic acids (see also T59904-11) and RNA polymerase II holoenzymes
CC can be used to treat diseases resulting from alteration or deletion
CC of the SRB gene, pref. by gene transfer technology. Probes based on
CC the genes can be used to detect SRB nucleic acids.
SQ Sequence 780 BP; 289 A; 145 C; 141 G; 205 T;

Query Match 1.2%; Score 36; DB 1; Length 780;
Best Local Similarity 100.0%; Pred. No. 6; 7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcataaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 729 TCATAAAAAAAAAAAAAAAAAAAAAA 764

RESULT 30
N92416
ID N92416 standard; cDNA; 794 BP.
AC N92416;
DT 15-MAR-1992 (first entry)
DE Sequence of the S3 cDNA clone.
KM Self-incompatible plant; S-gene; S-protein; ss.
OS Nicotiana glauca.
PN BP-343947-R.
PD 29-NOV-1989.
PR 24-MAY-1989; 305233.
PR 29-OCT-1985; US-792435.
PR 21-APR-1986; US-854139.
PR 25-MAY-1988; US-198781.
PA (LUBR) LOBRIZOL GENETICS I.
PI Clarke AE, Mau ST, Anderson MA, Cornish E, Tregear GW,
PI Crawford RJ, Miall HD, Bernatzky R;
DR WPI: 89-34992/48.
DR P-PSDB: P93469.
PT S-gene of a gametophytic self-incompatible plant - used to
PT produce S-protein and for genetic manipulation to create
PT self-incompatible cultivars
PS Example; Table 4, Page 20-21; 26pp; English.
CC The difference at the 3' end between clones NA-2-2 and NA-2-1
CC (see N92416) is near full-length but during sub-cloning a short
CC EcoRI fragment at the 5' end of the clone was inadvertently deleted.
CC The inventors claim a method for it isolating and identifying a cDNA
CC clone of an S-gene of a gametophytic self-incompatible plant (GSIP).
SQ Sequence 794 BP; 309 A; 136 C; 140 G; 207 T;

RESULT 31
X04314
ID X04314 standard; DNA; 808 BP.
AC X04314;
DT 13-APR-1999 (first entry)
DE Human secreted protein gene 4 clone HOVA158.
KM Human; secreted protein; fusion protein; gene therapy; protein therapy;
KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KM developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KM inflammation; ischaemic shock; Alzheimer's disease; osteoarthritis; AIDS;
KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN W09856804-A1.
PD 17-DEC-1998.
PR 11-JUN-1998; U12125.
PR 02-OCT-1997; US-061060.
PR 13-JUN-1997; US-049547.
PR 13-JUN-1997; US-049548.
PR 13-JUN-1997; US-049549.
PR 13-JUN-1997; US-049550.
PR 13-JUN-1997; US-049606.
PR 13-JUN-1997; US-049607.
PR 13-JUN-1997; US-049608.
PR 13-JUN-1997; US-049609.
PR 13-JUN-1997; US-049610.
PR 13-JUN-1997; US-049611.
PR 13-JUN-1997; US-050566.
PR 13-JUN-1997; US-050901.
PR 13-JUN-1997; US-052989.
PR 08-JUL-1997; US-051919.
PR 18-AUG-1997; US-055984.
PR 12-SEP-1997; US-058665.
PR 12-SEP-1997; US-058666.
PR 12-SEP-1997; US-058669.
PR 12-SEP-1997; US-058750.
PR 12-SEP-1997; US-058971.
PR 12-SEP-1997; US-058972.
PR 12-SEP-1997; US-058975.
PR 02-OCT-1997; US-060834.
PR 02-OCT-1997; US-060841.
PR 02-OCT-1997; US-060844.
PR 02-OCT-1997; US-060865.
PR 02-OCT-1997; US-061059.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Ferrile AM, Feng P, Greene JM, Lafleur DM,
PI Moore PA, Nl J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P,
PI Yu GL.
DR WPI: 99-080881/07.
DR P-PSDB: W78129.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1; Page 169-170; 380pp; English.
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number, and the clone it is derived from, are
CC detailed in the descriptor line. The gene can be used to generate fusion
CC proteins by linking to the gene to a human immunoglobulin Fc portion
CC (e.g. X04302) for increasing the stability of the fused protein as
CC compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic acid
CC sequences: X04311-X04410; amino acid sequences W78126-W78225) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on
CC which tissues they are most highly expressed in (see X04311 for described
CC uses).

SQ Sequence 808 BP; 268 A; 165 C; 148 G; 225 T;

Query Match 1.2%; Score 36; DB 1; Length 808;

Best Local Similarity 100.0%; Pred. No. 6.5e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 756 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 791

RESULT 32

X51754

ID X51754 standard; DNA; 886 BP.

AC X51754;

DT 17-JUN-1999 (first entry)

DE DNA encoding a human secreted protein.

KW Human secreted protein; cancer; immune disorder; infection;

KW inflammatory disorder; skin disorder; tumour; atherosclerosis;

KW periphereal neuropathy; trauma; spinal cord injury; allergy;

KW hematopoietic disorder; skeletal disorder; neurological disorder;

KW arthritic disorder; asthma; immunodeficiency disease; AIDS;

KW transplant rejection; ss.

OS Homo sapiens.

PN W09911293-A1.

PD 11-MAR-1999.

PF 03-SEP-1998; U18360.

PR 12-SEP-1997; US-058974.

PR 05-SEP-1997; US-057626.

PR 05-SEP-1997; US-057663.

PR 05-SEP-1997; US-057669.

PR 12-SEP-1997; US-058666.

PR 12-SEP-1997; US-058667.

PR 12-SEP-1997; US-058973.

PA (HUMA-) HUMAN GENE SCIT INC.

PI Brewer LA, Ebner R, Lafleur DW, Moore PA, Olsen HS,

PI Rosen GA, Ruben SM, Shi Y;

DR WPI; 99-204988/17.

DR P-PSDB; Y12967.

PT New isolated human genes and the secreted polypeptides they encode

PT - useful for diagnosis and treatment of e.g. neurological disorders,

PT tumours, immune disorders, inflammation or haematological disorders,

PS Claim 1; Page 180; 215pp; English.

CC X51701-55 encode human secreted proteins. The polynucleotides and

CC their corresponding secreted polypeptides are useful for preventing,

CC treating or ameliorating medical conditions, e.g. by protein or gene

CC therapy. Pathological conditions can also be diagnosed by determining

CC the amount of the new polypeptides in a sample or by determining the

CC presence of mutations in the new polynucleotides. Specific uses are

CC described for each polynucleotide, based on which tissues they are

CC most highly expressed in, and include developing products for the

CC diagnosis or treatment of cancer, immune disorders, infection,

CC inflammatory disorders, skin disorders, tumours, atherosclerosis,

CC restenosis, autoimmune disorders, Alzheimer's disease, peripheral

CC neuropathies, trauma, spinal cord injuries, allergy, hematopoietic

CC disorders, skeletal disorders, neurological disorders, arthritic

CC disorders, asthma, immunodeficiency diseases, AIDS and transplant

CC rejection. The polypeptides are also useful for identifying their

CC binding partners.

SQ Sequence 886 BP; 288 A; 163 C; 174 G; 261 T;

Query Match 1.2%; Score 36; DB 1; Length 886;

Best Local Similarity 100.0%; Pred. No. 6.5e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 801 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 836

RESULT 33

ID Q92529 standard; cDNA; 893 BP.

AC Q92529;

DT 09-DEC-1995 (first entry)

DE P. communis (pear) arabinogalactan protein from cDNA clone PCAGP9.

KW Arabinogalactan protein; AGP; gum; thickening agent; emulsifier;

KW adhesive; ss.

OS Pyrus communis.

FN Key Location/Qualifiers

FT cds 70..507

FT //tag= a

PN W09515377-A.

PD 08-JUN-1995.

PF 01-DEC-1994; AU0744.

PR 03-DEC-1993; US-161944.

PR 18-JUL-1994; US-276452.

PA (ALBR) ALBRIGHT & WILSON AUSTRALIA.

PI Bactic A, Chen C, Clarke AE, Du H, Gane AM, Mau S;

PI WPI; 95-215258/28.

DR P-PSDB; R75511.

PT Recombinant plant arabinogalactan protein and cloned DNA encoding it

PT useful as an emulsifying agent, adhesive agent or a lubricating agent.

PS Claim 22; Figure 5A, 142pp; English.

CC Total native arabinogalactan protein (AGP) was isolated from

CC P. communis. It was deglycosylated using anhydrous HF.

CC The total deglycosylated AGP sample was then reduced and

CC carboxymethylated. It was then separated by HPLC. The fractions

CC were subjected to thermolysis digestion and then separated.

CC Individual peaks from thermolysis digested were sequenced.

CC The sequence R75536 was selected as a template for the isolation

CC or the corresp. AGP gene. An antisense RNA probe was synthesised

CC from the PCR template and used to screen a cDNA library prepd.

CC from pear cell suspension culture. Three cDNA clones were isolated

CC and sequenced. The sequence of the longest cDNA clone PCAGP9 is

CC shown in Q92529/R75511.

SQ Sequence 893 BP; 208 A; 248 C; 162 G; 275 T;

Query Match 1.2%; Score 36; DB 1; Length 893;

Best Local Similarity 100.0%; Pred. No. 6.5e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 855 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 890

RESULT 34

ID V99723 standard; cDNA; 894 BP.

AC V99723;

DT 26-APR-1999 (first entry)

DE Human adult retina secreted protein bx200_13 cDNA.

KW Secreted protein; human; retina; bx200_13; ds.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 245..523

FT //tag= a

PN W09856909-A2.

PD 17-DEC-1998.

PF 08-JUN-1998; U11822.

PR 05-JUN-1998; US-092722.

PR 11-JUN-1997; US-873218.

PA (GENM) GENETICS INST INC.

PI Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavallie ER,

PI McCoy JM, Racie LA, Spaulding V, Treacy M;

PI WPI; 99-080899/07.

DR P-PSDB; W95346.

PT New polynucleotides encoding secreted human proteins - derived from

PT human foetal brain, adult testes, foetal kidney, adult thyroid or

PT adult retina cDNA libraries

PS Claim 15; Page 74; 113pp; English.

CC This is the nucleotide sequence of cDNA clone bk200_13, which
 CC includes an open reading frame for a 92-amino acid polypeptide
 CC (see W95346). The clone was isolated from a human adult retina
 CC cDNA library using methods which are selective for cDNAs encoding
 CC secreted proteins, or was identified as encoding a secreted or
 CC transmembrane protein on the basis of computer analysis of the
 CC amino acid sequence of the encoding protein. Database searches
 CC indicate some sequence similarity to known sequences. The
 CC inventory provides cDNA clones (see V99721-33) from human adult
 CC thyroid, adult retina, adult testis, foetal kidney and foetal brain
 CC that encode novel secreted proteins (see W95344-53). Each clone is
 CC individually available from deposit clone ATCC 98451 (see also
 CC V99734-43). The isolated polynucleotides (PNS) and proteins are
 CC predicted to have activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans
 CC and animals, although no supporting data is given. Suggested
 CC activities include nutritional, cytokine, cell proliferation or
 CC differentiation, immune stimulating (e.g. as vaccines) or immune
 CC suppressing, haematopoiesis regulating, tissue growth,
 CC activin/inhibin, chemotactic/chemokinetic, haemostatic,
 CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activities. The PNS are
 CC also stated to be useful for gene therapy. 184 G; 319 T;
 SQ Sequence 894 BP; 274 A; 117 C; 184 G; 319 T;

Query Match 1.2%; Score 36; DB 1; Length 894;
 Best Local Similarity 100.0%; Pred. No. 6.5e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaa 3117
 |||||
 Db 852 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 887

RESULT 35
 X27330
 ID X27330 standard; DNA; 921 BP.
 AC X27330;
 DT 11-JUN-1999 (first entry)
 DE Human secreted protein gene 20 clone HMKAH10.
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 PN W09902546-A1.
 PD 21-JAN-1999.
 PF 07-JUL-1998; U13684.
 PR 12-SEP-1997; US-058785.
 PR 08-JUL-1997; US-051916.
 PR 08-JUL-1997; US-051918.
 PR 08-JUL-1997; US-051919.
 PR 08-JUL-1997; US-051920.
 PR 08-JUL-1997; US-051925.
 PR 08-JUL-1997; US-051926.
 PR 08-JUL-1997; US-051928.
 PR 08-JUL-1997; US-051929.
 PR 08-JUL-1997; US-051930.
 PR 08-JUL-1997; US-051931.
 PR 08-JUL-1997; US-051932.
 PR 08-JUL-1997; US-052732.
 PR 08-JUL-1997; US-052733.
 PR 08-JUL-1997; US-052735.
 PR 08-JUL-1997; US-052795.
 PR 08-JUL-1997; US-052803.
 PR 18-AUG-1997; US-055684.
 PR 18-AUG-1997; US-055722.
 PR 18-AUG-1997; US-055723.

PR 18-AUG-1997; US-055947.
 PR 18-AUG-1997; US-055948.
 PR 18-AUG-1997; US-055949.
 PR 18-AUG-1997; US-055950.
 PR 18-AUG-1997; US-055953.
 PR 18-AUG-1997; US-055954.
 PR 18-AUG-1997; US-055964.
 PR 18-AUG-1997; US-055984.
 PR 18-AUG-1997; US-056360.
 PR 12-SEP-1997; US-058660.
 PR 12-SEP-1997; US-058661.
 PR 12-SEP-1997; US-058664.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Bresler LA, Ebner R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA,
 PI Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Zeng Z;
 DR WPI; 99-120770/10.
 DR P-PSDB; Y02669.
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 PS Claim 1: Page 256; 464pp; English.
 CC This sequence represents a nucleic acid molecule which encodes a secreted
 CC human protein. The gene number, and the clone it is derived from, are
 CC detailed in the descriptor line. The gene can be used to generate fusion
 CC proteins by linking to the gene to a human immunoglobulin Fc portion
 CC (e.g. X27302) for increasing the stability of the fused protein as
 CC compared to the human protein only.
 CC The invention relates to 123 novel genes and their fragments (nucleic
 CC acid sequences: X27311-X27449; amino acid sequences Y02650-Y02788) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 123 polynucleotides, based on
 CC which tissues they are most highly expressed in (see X27311 for described
 CC uses).
 SQ Sequence 921 BP; 258 A; 187 C; 195 G; 281 T;

Query Match 1.2%; Score 36; DB 1; Length 921;
 Best Local Similarity 100.0%; Pred. No. 6.5e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaa 3117
 |||||
 Db 875 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 910

RESULT 36
 X51728
 ID X51728 standard; DNA; 921 BP.
 AC X51728;
 DT 17-JUN-1999 (first entry)
 DE DNA encoding a human secreted protein.
 KW Human secreted protein; cancer; immune disorder; infection;
 KW inflammatory disorder; skin disorder; tumour; atherosclerosis;
 KW restenosis; autoimmune disorder; Alzheimer's disease;
 KW peripheral neuropathy; trauma; spinal cord injury; allergy;
 KW hematopoietic disorder; skeletal disorder; neurological disorder;
 KW arthritic disorder; asthma; immunodeficiency disease; AIDS;
 KW transplant rejection; ss.
 OS Homo sapiens.
 PN W09911293-A1.
 PD 11-MAR-1999.
 PF 03-SEP-1998; U18360.
 PR 12-SEP-1997; US-058974.
 PR 05-SEP-1997; US-057626.
 PR 05-SEP-1997; US-057663.
 PR 05-SEP-1997; US-057669.
 PR 12-SEP-1997; US-058666.
 PR 12-SEP-1997; US-058667.
 PR 12-SEP-1997; US-058973.
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Brewer LA, Ebner R, Lafleur DW, Moore PA, Olsen HS,
 PI Rosen GA, Ruben SM, Shi Y;
 DR WPI: 99-204988/17.
 DR P-PSDB; Y12941.
 PT New isolated human genes and the secreted polypeptides they encode
 PT - useful for diagnosis and treatment of e.g. neurological disorders,
 PT tumours, immune disorders, inflammation or haematological disorders
 PS Claim 1; Page 169; 215pp; English.
 CC X51701-55 encode human secreted proteins. The polynucleotides and
 CC their corresponding secreted polypeptides are useful for preventing,
 CC treating or ameliorating medical conditions, e.g. by protein or gene
 CC therapy. Pathological conditions can also be diagnosed by determining
 CC the amount of the new polypeptides in a sample or by determining the
 CC presence of mutations in the new polynucleotides. Specific uses are
 CC described for each polynucleotide, based on which tissues they are
 CC most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, immune disorders, infection,
 CC inflammatory disorders, skin disorders, tumours, atherosclerosis,
 CC restenosis, autoimmune disorders, Alzheimer's disease, peripheral
 CC neuropathies, trauma, spinal cord injuries, allergy, hematopoietic
 CC disorders, skeletal disorders, neurological disorders, arthritic
 CC disorders, asthma, immunodeficiency diseases, AIDS and transplant
 CC rejection. The polypeptides are also useful for identifying their
 CC binding partners.
 SQ Sequence 921 BP; 250 A; 218 C; 174 G; 279 T;

Query Match 1.2%; Score 36; DB 1; Length 921;
 Best Local Similarity 100.0%; Pred. No. 6.5e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 DB 877 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 912

RESULT 37

ID Q73760 standard; DNA; 971 BP.
 AC Q73760; 995 (first entry)
 DE RP-8 Programmed cell death polypeptide gene.
 KW Programmed cell death; RP-8; PCD; cancer cells; immune cells;
 KW autoimmune disease; Alzheimer's disease; Huntington's disease;
 KW Parkinson's disease; ss.
 OS Rattus rattus.
 FH Key Location/Qualifiers
 FT cds 1..864
 FT /*tag= a
 FT /product= RP-8 Programmed cell death polypeptide.
 PN US5360893-A.
 PD 01-NOV-1994.
 PF 20-JUL-1992; 915934.
 PR 20-JUL-1992; US-915934.
 PA (COLS) UNIV COLORADO FOUND INC.
 PI Cohen JJ, Hahn WE, Owens GP;
 DR WPI: 94-349500/43.
 DR P-PSDB; R60866.
 PT New polypeptides eliciting programmed cell death - used to
 PT develop prods. for detecting, activating or preventing programmed
 PT cell death
 PS Claim 1: Columns 21-24; 23pp; English.
 CC The nucleic acid, RP-8 programmed cell death (PCD) polypeptide and
 CC antibodies directed against the polypeptide can be used to develop
 CC products for use in diagnosis and therapy. The products can be used
 CC for detecting PCD, for activating PCD in unwanted cells such as
 CC cancer cells and immune cells linked to autoimmune diseases or for
 CC preventing unwanted cell death in degenerative disorders such as
 CC Alzheimer's disease, Parkinson's disease and Huntington's disease.
 SQ Sequence 971 BP; 312 A; 204 C; 225 G; 230 T;

Query Match 1.2%; Score 36; DB 1; Length 971;

Best Local Similarity 100.0%; Pred. No. 6.4e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 DB 910 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 945

RESULT 38

ID T35704 standard; DNA; 971 BP.
 AC T35704;
 DE 07-OCT-1996 (first entry)
 KW Coding sequence for programmed mammalian cell death protein RP-8.
 KW Programmed mammalian cell death; PMCD; RP-8; human disease; cancer; RP-2;
 KW autoimmune disease; therapy; chemotherapy; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT cds 1..864
 FT /*tag= a
 FT /product= RP-8
 PN US5527682-A.
 PD 18-JUN-1996.
 PF 20-JUL-1992; 915934.
 PR 20-JUL-1992; US-915934.
 PR 19-OCT-1994; US-325743.
 PA (COLS) UNIV COLORADO.
 PI Cohen JJ, Hahn WE, Owens GP;
 DR WPI: 96-299855/30.
 DR P-PSDB; W00580.
 PT DNA encoding polypeptide(s) eliciting programmed mammalian cell
 PT death - useful to detect programmed cell death in various cells,
 PT and to treat and diagnose human diseases, e.g. cancer
 PS Claim 1: Column 19-22; 23pp; English.
 CC T35704 and T35705 represent the coding sequences for two proteins which
 CC elicit programmed mammalian cell death (PMCD). This encodes a PMCD
 CC protein designated RP-8. These sequences can be used in the treatment
 CC and diagnosis of human diseases, such as cancer, and to selectively
 CC eliminate immune cells linked to autoimmune diseases. These genes can
 CC also be used to monitor the extent of cell death associated with specific
 CC diseases, thereby allowing a course of treatment to be decided. A
 CC selective cancer therapy that works by selective activation of the body's
 CC own RP-8 and RP-2 genes will lessen, or eliminate the need for
 CC chemotherapy.
 SQ Sequence 971 BP; 312 A; 204 C; 225 G; 230 T;

Query Match 1.2%; Score 36; DB 1; Length 971;
 Best Local Similarity 100.0%; Pred. No. 6.4e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 DB 910 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 945

RESULT 39

ID X51748 standard; DNA; 990 BP.
 AC X51748;
 DT 17-JUN-1999 (first entry)
 DE DNA encoding a human secreted protein.
 KW Human secreted protein; cancer; immune disorder; infection;
 KW inflammatory disorder; skin disorder; tumour; atherosclerosis;
 KW restenosis; autoimmune disorder; Alzheimer's disease;
 KW peripheral neuropathy; trauma; spinal cord injury; allergy;
 KW hematopoietic disorder; skeletal disorder; neurological disorder;
 KW arthritic disorder; asthma; immunodeficiency disease; AIDS;
 KW transplant rejection; ss.
 OS Homo sapiens.
 PN W09911293-AI.
 PD 11-MAR-1999.
 PF 03-SEP-1998; U18360.

CC expression vector, a cultured cell carrying the vector, and an
 CC oligonucleotide probe or primer (see V45446). The cultured cell
 CC may be used for the production of protein (Claimed).
 CC ZS1G-35 may be used in therapeutics (claimed), for the regulation of
 CC acute and chronic inflammatory disease conditions, lymphocyte
 CC migration and ischaemia/reperfusion injury. The probe can be used
 CC to detect a genetic abnormality in a patient (claimed) by comparing
 CC the hybridisation of a genetic sample to the probe to produce a
 CC first reaction product, and then comparing this product to a
 CC control.
 CC Sequence 1029 BP; 241 A; 303 C; 261 G; 224 T;

Query Match 1.2%; Score 36; DB 1; Length 1029;

Best Local Similarity 100.0%; Pred. No. 6.4e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3117
 |||||AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA|||
 DB 977 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1012

RESULT 42

ID X30150 standard; DNA; 1037 BP.

AC X30150;

DT 18-JUN-1999 (first entry)

DE Human secreted protein gene 6.

KW Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; ss.

OS Homo sapiens.

PN W0910363-A1.

PD 04-MAR-1999.

PF 27-AUG-1998; U17709.

PR 29-AUG-1997; US-056271.

PR 29-AUG-1997; US-056073.

PR 29-AUG-1997; US-056247.

PR 29-AUG-1997; US-056270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Fan P, Kyaw H, Rosen CA, Ruben SM, Wei YF;

DR WPI: 99-190585/16.

DR F-PSDB; Y04298.

PT New isolated human genes and the secreted polypeptides they encode
 PT - useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders

PS Claim 1; Page 135; 170pp; English.

CC X30145 to X30173 represent 29 isolated human secreted protein genes.
 CC Y04293 to Y04321 represent the secreted proteins encoded by the 29 human
 CC genes. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 29 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.

CC The sequences given in X30174 to X30182 and Y04322 to Y04334 are used
 CC in the exemplification of the present invention.

CC Sequence 1037 BP; 308 A; 198 C; 224 G; 306 T;

SQ

Query Match 1.2%; Score 36; DB 1; Length 1037;

Best Local Similarity 100.0%; Pred. No. 6.4e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3117

|||||AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA|||

DB 995 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1030

RESULT 43

QI0377 QI0377 standard; DNA; 1046 BP.

ID QI0377;

AC QI0377;

DT 12-APR-1991 (first entry)

DE Plasmid pMG3C9 used to isolate style-stigma specific gene STG3C9.

KW Hybrid vigour; pollination; ss.

OS Nictiana tabacum "Petit Havana" SRL.

FH Key Location/Qualifiers

FT cds 3..803

FT /*tag= a

PN EP-412006-A.

PD 06-FEB-1991.

PF 31-JUL-1990; 402196.

PR 04-AUG-1989; EP-402224.

PR 31-JUL-1990; EP-402196.

PA (PLAN-) PLANT GENETIC SYST.

PI De Greef W, Van Emmelo J, De Oliveira DE, De Souza MH;

PI Van Montagu M;

DR WPI: 91-038878/06.

DR P-PSDB; R10535.

PT Modified plant with transformed nuclear genome - obt'd. by using
 PT foreign DNA encoding prod. which disrupts metabolism, functioning
 PT and/or development of plant.

PS Disclosure: Fig 2B; 42pp; English.

CC The clone was produced by subcloning a style-stigma specific cDNA,
 CC 3C9, obt'd. from Prof. Goldberg of UCLA, into pGEM1, to produce
 CC pMG3C9. The clone was found to be stigma-style specific in
 CC Northern analysis. A probe from the clone was used to isolate

CC the corresp. genomic sequence which is specifically expressed in
 CC style-stigma tissues of the female organ of tobacco. The corresp.
 CC clone, pSTG3C9, contains STMG3C9, an "STMG-type" gene. This and
 CC other similar genes can be used to produce new hybrid plants or
 CC seeds having a combination of desirable traits, and showing hybrid
 CC vigour. A female-sterile, male-fertile plant can be produced which
 CC favours cross-pollination.

CC See also Q10374-76.

CC Sequence 1046 BP; 360 A; 254 C; 166 G; 266 T;

SQ

Query Match 1.2%; Score 36; DB 1; Length 1046;

Best Local Similarity 100.0%; Pred. No. 6.4e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3117

|||||AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA|||

DB 964 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 999

RESULT 44

ID X37464

AC X37464 standard; cDNA; 1060 BP.

DT 06-JUL-1999 (first entry)

DE Human secreted protein cDNA fragment containing gene 14.

KW Human; secreted protein; treatment; prevention; protein therapy; AIDS;

KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;

KW developmental abnormality; fetal deficiency; blood disorder; leukemia;

KW immune system disease; autoimmune disease; hepatic disease; lymphoma;

KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;

KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;

KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;

KW arthritis; malignancy; digestive; endocrine; infection; ss.

OS Homo sapiens.

PN W0918208-A1.

PD 15-APR-1999.

```
PF 01-OCT-1998; U20775.
PR 02-OCT-1997; US-060884.
PR 02-OCT-1997; US-060833.
PR 02-OCT-1997; US-060836.
PR 02-OCT-1997; US-060837.
PR 02-OCT-1997; US-060838.
PR 02-OCT-1997; US-060839.
PR 02-OCT-1997; US-060843.
PR 02-OCT-1997; US-060862.
PR 02-OCT-1997; US-060866.
PR 02-OCT-1997; US-060874.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Carter KC, Duan DR, Endress GA, Feng P, Ferrie AM,
PI Florence KA, Greene JM, Janat F, Lafleur DW, Ni J,
PI Rosen CA, Ruben SM, Shi Y, Young P, Yu G;
DR WPI: 99-264022/22.
DR P-PSDB; Y07865.
PT New isolated human genes and the secreted polypeptides they encode
PS Claim 1a; Page 235-236; 368pp; English.
CC This invention describes novel isolated human genes and the secreted
CC proteins they encode. The products of the invention are useful for
CC preventing, treating or ameliorating medical conditions, e.g. by protein
CC or gene therapy. Also pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 101 polynucleotides,
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours,
CC neurodegenerative disorders, developmental abnormalities and fetal
CC deficiencies, blood disorders, leukemias, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,
CC allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate
CC disease, skeletal or cardiac muscle disorders, pulmonary disorders,
CC transplant rejection, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, digestive/endocrine disorders,
CC infections and AIDS. The human secreted proteins of the invention are
CC represented in Y07852-Y07993 and the encoding nucleic acids are
CC represented in X37451-X37552.
SQ Sequence 1060 BP; 328 A; 205 C; 192 G; 326 T;

Query Match 1.2%; Score 36; DB 1; Length 1060;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 958 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 993

RESULT 45
X37460
ID X37460 standard; cDNA; 1089 BP.
AC X37460;
DT 06-JUL-1999 (first entry)
DE Human secreted protein cDNA fragment containing gene 10.
KW Human; secreted protein; treatment; prevention; protein therapy; AIDS;
KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW immune system disease; autoimmune disease; hepatic disease; lymphoma;
KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;
KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;
KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;
KW arthritis; malignancy; digestive; endocrine; infection; ss.
OS Homo sapiens.
PN WO9918208-A1.
PD 15-APR-1999.
PF 01-OCT-1998; U20775.
PR 02-OCT-1997; US-060884.
PR 02-OCT-1997; US-060833.
PR 02-OCT-1997; US-060836.
PR 02-OCT-1997; US-060837.
PR 02-OCT-1997; US-060838.
```

```
PR 02-OCT-1997; US-060839.
PR 02-OCT-1997; US-060843.
PR 02-OCT-1997; US-060862.
PR 02-OCT-1997; US-060866.
PR 02-OCT-1997; US-060874.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Carter KC, Duan DR, Endress GA, Feng P, Ferrie AM,
PI Florence KA, Greene JM, Janat F, Lafleur DW, Ni J,
PI Rosen CA, Ruben SM, Shi Y, Young P, Yu G;
DR WPI: 99-264022/22.
DR P-PSDB; Y07861.
PT New isolated human genes and the secreted polypeptides they encode
PS Claim 1a; Page 232-233; 368pp; English.
CC This invention describes novel isolated human genes and the secreted
CC proteins they encode. The products of the invention are useful for
CC preventing, treating or ameliorating medical conditions, e.g. by protein
CC or gene therapy. Also pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 101 polynucleotides,
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours,
CC neurodegenerative disorders, developmental abnormalities and fetal
CC deficiencies, blood disorders, leukemias, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,
CC allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate
CC disease, skeletal or cardiac muscle disorders, pulmonary disorders,
CC transplant rejection, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, digestive/endocrine disorders,
CC infections and AIDS. The human secreted proteins of the invention are
CC represented in Y07852-Y07993 and the encoding nucleic acids are
CC represented in X37451-X37552.
SQ Sequence 1089 BP; 367 A; 202 C; 240 G; 279 T;

Query Match 1.2%; Score 36; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 1042 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1077

Search completed: September 11, 2000, 00:14:14
Job time: 7147 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:40:42 ; Search time 1087.01 Seconds
(without alignments)
12645.178 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtgaccacgcgctcg.....aaaaaaaaaaaaaaaaaaaaa 3117

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5247842 seqs, 2204914090 residues

Word size : 0
Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_gss1:*
94: gb_gss2:*
95: gb_gss3:*
96: gb_gss4:*
97: em_gss1:*
98: em_gss2:*
99: em_gss3:*
100: em_gss4:*
101: gb_gss5:*
102: gb_gss6:*
103: gb_gss7:*
104: gb_gss8:*
105: gb_gss9:*
106: em_gss5:*
107: em_gss6:*
108: em_gss7:*
109: em_gss8:*
110: em_gss9:*
111: em_gss10:*
112: em_gss11:*
113: gb_gss10:*
114: gb_gss11:*
115: em_gss12:*
116: gb_gss12:*

117: gb_gss13:*
118: gb_gss14:*
119: gb_gss15:*
120: gb_gss16:*
121: gb_gss17:*
122: gb_gss18:*
123: gb_gss19:*
124: em_gss13:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	576	18.5	657	AI013917	AI013917 EST208592
C 2	65	2.1	387	AA925917	AA925917 DT-R-A1-e
C 3	61	2.0	496	AI891958	AI891958 ul57a09.y
C 4	52	1.7	349	AW325627	AW325627 17814 MAR
C 5	45	1.4	442	AA980401	AA980401 ua52911.r
C 6	41	1.3	157	AL040100	AL040100 DKFZp434B
C 7	41	1.3	157	AW278153	AW278153 sf40808.y
C 8	41	1.3	240	AU073712	AU073712 AU073712
C 9	41	1.3	262	AU065863	AU065863 AU065863
C 10	41	1.3	1367	AF034791	AF034791 AF034791
C 11	40	1.3	210	AW313563	AW313563 8840 MARC
C 12	40	1.3	224	AW657118	AW657118 109764 MA
C 13	40	1.3	275	AI896601	AI896601 C98601 Rice
C 14	40	1.3	449	AI253246	AI253246 qz39e12.x
C 15	39	1.3	79	AI320733	AI320733 C9f03nm.f
C 16	39	1.3	139	AI946912	AI946912 ps2h08.y
C 17	39	1.3	161	AI540624	AI540624 tn88609.x
C 18	39	1.3	167	AA231085	AA231085 mw11e11.r
C 19	39	1.3	182	AI630474	AI630474 ad11a06.y
C 20	39	1.3	190	73 AW477617	73 AW477617 15122 MAR
C 21	39	1.3	207	AW172002	AW172002 618047C10
C 22	39	1.3	229	71 AW357288	71 AW357288 40265 MAR
C 23	39	1.3	285	AA832425	AA832425 cc99d12.s
C 24	39	1.3	281	AW056177	AW056177 660005C01
C 25	39	1.3	307	73 AW479416	73 AW479416 25195 MAR
C 26	39	1.3	318	AA237390	AA237390 mx18b02.r
C 27	39	1.3	327	80 AU082521	80 AU082521 AU082521
C 28	39	1.3	337	30 AA773444	30 AA773444 ad58d05.s
C 29	39	1.3	340	AA595839	AA595839 nn05c01.s
C 30	39	1.3	341	71 AW394641	71 AW394641 sh06a07.y
C 31	39	1.3	376	40 AI529904	40 AI529904 ul83h01.y
C 32	39	1.3	380	81 D22652	81 D22652 RICC0739B R
C 33	39	1.3	424	44 AI829162	44 AI829162 wk76a05.x
C 34	39	1.3	427	34 AI050855	34 AI050855 oy47a05.x
C 35	39	1.3	444	28 AA610049	28 AA610049 af18h01.s
C 36	39	1.3	450	30 AA731682	30 AA731682 nw58e10.s
C 37	39	1.3	475	70 AW249071	70 AW249071 2820891.5
C 38	39	1.3	482	34 AI018664	34 AI018664 ov65e05.s
C 39	39	1.3	499	25 AA393950	25 AA393950 zt178a10.r
C 40	39	1.3	506	31 AA808304	31 AA808304 oc41f12.s
C 41	39	1.3	510	31 AA810286	31 AA810286 oc41h11.s
C 42	39	1.3	519	37 AI2411203	37 AI2411203 qj96g10.x
C 43	39	1.3	520	47 AT000142	47 AT000142 AT000142
C 44	39	1.3	523	39 AI397624	39 AI397624 NCSC5611T
C 45	39	1.3	523	74 AW566146	74 AW566146 660082F05
C 46	39	1.3	738	36 AI187484	36 AI187484 EST273 Ma
C 47	38	1.2	96	28 AA607306	28 AA607306 vm96f10.r
C 48	38	1.2	115	42 AI619093	42 AI619093 486073C06
C 49	38	1.2	145	31 AA829401	31 AA829401 od06406.s
C 50	38	1.2	166	41 AI567226	41 AI567226 tp51f02.x
C 51	38	1.2	169	64 AW164849	64 AW164849 se78a05.y
C 52	38	1.2	179	73 AW477009	73 AW477009 ga40604.y
C 53	38	1.2	188	23 AA280902	23 AA280902 zsz97f10.s
C 54	38	1.2	188	27 AA497258	27 AA497258 fa04e08.s
C 55	38	1.2	189	79 AW656645	79 AW656645 109003 MA

38	1.2	193	63	AW100662	AW100662
38	1.2	203	73	AA485903	AA485903
38	1.2	208	32	AA869427	AA869427 vc07h10.r
38	1.2	234	74	AW583947	AW583947 ia07d01.x
38	1.2	239	63	AW102081	AW102081 sg82908.y
38	1.2	242	70	AW278164	AW278164 sf40e07.y
38	1.2	244	43	AI695228	AI695228 wa01g09.x
38	1.2	245	30	AA738321	AA738321 nx16h11.s
38	1.2	246	35	AI092654	AI092654 qa80b07.x
38	1.2	264	48	AU082927	AU082927
38	1.2	272	30	AA711214	AA711214 vt70b06.r
38	1.2	278	40	AI491852	AI491852 to01e04.x
38	1.2	280	31	AA830261	AA830261 oc45f08.s
38	1.2	284	81	C97620	C97620 C97620 Rice
38	1.2	285	41	AI563599	AI563599 vx91b04.x
38	1.2	289	47	AU033170	AU033170
38	1.2	297	47	AU033074	AU033074
38	1.2	302	47	AU030838	AU030838
38	1.2	304	70	AW311758	AW311758 5631 MARC
38	1.2	318	22	AA154878	AA154878 mx32e05.r
38	1.2	322	31	AA808311	AA808311 oc41g07.s
38	1.2	322	64	AI166640	AI166640 xg54g08.x
38	1.2	323	40	AI499147	AI499147 to05e09.x
38	1.2	325	62	AW021091	AW021091 df18f09.y
38	1.2	329	91	W79740	W79740 zd86g04.r1
38	1.2	337	81	C98639	C98639 C98639 Rice
38	1.2	338	80	AW733654	AW733654 sr83c12.y
38	1.2	341	30	AA729009	AA729009 nw25g12.s
38	1.2	341	81	C93971	C93971 C93971 D1ct
38	1.2	343	48	AU075394	AU075394
38	1.2	346	80	AW734302	AW734302 sk81d04.y
38	1.2	352	81	C99307	C99307 C99307 Rice
38	1.2	353	70	AW307000	AW307000 sf51e06.y
38	1.2	356	23	AA277361	AA277361 va78g10.r
38	1.2	356	48	AU0771018	AU0771018
38	1.2	357	31	AA837811	AA837811 oe39g07.s
38	1.2	363	32	AA871182	AA871182 vg32b04.r
38	1.2	369	42	AI681953	AI681953 tp45a07.x
38	1.2	372	44	AI822331	AI822331 LO-776T3
38	1.2	373	28	AA609483	AA609483 af14b02.s
38	1.2	384	22	AA198948	AA198948 mv39g03.r
38	1.2	390	27	AA495424	AA495424 fa03b01.s
38	1.2	402	70	AW311095	AW311095 sg32g08.y
38	1.2	411	21	AA129588	AA129588 z11d09.s
38	1.2	414	22	AA217736	AA217736 mv56g03.r
38	1.2	417	21	AA127454	AA127454 zn92f10.s
38	1.2	423	31	AA831663	AA831663 oc83g12.s
38	1.2	431	86	N26206	N26206 yw53a09.s1
38	1.2	435	23	AA281501	AA281501 zt08g04.s
38	1.2	448	47	AU029299	AU029299
38	1.2	450	48	AU039070	AU039070
38	1.2	467	21	AA131365	AA131365 zo08b06.s
38	1.2	483	31	AA810223	AA810223 od14a09.s
38	1.2	484	46	AI973520	AI973520 496027F11
38	1.2	486	31	AA836268	AA836268 od17e05.s
38	1.2	490	31	AA832000	AA832000 oc97c11.s
38	1.2	497	28	AA603342	AA603342 np29d07.s
38	1.2	510	31	AA827614	AA827614 of04h03.s
38	1.2	521	29	AA678835	AA678835 ah09d04.s
38	1.2	533	27	AA490301	AA490301 va51a03.s
38	1.2	548	30	AA710038	AA710038 zt38g03.r
38	1.2	548	37	AI257985	AI257985 LP06506.s
38	1.2	548	48	AU058000	AU058000
38	1.2	548	69	AW186926	AW186926 BNLGHI738
38	1.2	552	31	AA825404	AA825404 oc57h10.s
38	1.2	563	48	AU056286	AU056286
38	1.2	583	39	AI398912	AI398912 NCW06bD77
38	1.2	592	48	AU086008	AU086008
38	1.2	615	21	AA108889	AA108889 ma15eb11.r
38	1.2	628	63	AW090702	AW090702 xc95b11.x
38	1.2	637	69	AW189241	AW189241 x103b12.x
38	1.2	653	31	AA837091	AA837091 od18h01.s
38	1.2	703	47	AU030679	AU030679

129	38	1.2	706	81	C97853	C97853 C97853 Rice	c 202	37	1.2	256	29	AA643982	AA643982 nj40h09.s
130	37	1.2	50	28	AA590944	AA590944 vm25f02.r	203	37	1.2	256	47	AU029273	AU029273 AU029273
131	37	1.2	56	79	AW632959	AW632959 bl02c02.x	204	37	1.2	258	72	AW424013	AW424013 sh59q12.y
132	37	1.2	69	28	AA607058	AA607058 vm95e07.r	205	37	1.2	262	63	AW099006	AW099006 sd33a03.y
133	37	1.2	69	31	AA795179	AA795179 vq94g07.r	206	37	1.2	263	72	AW432647	AW432647 sh83e05.y
134	37	1.2	77	89	R99209	R99209 yq64e12.r1	207	37	1.2	264	47	AU029317	AU029317 AU029317
135	37	1.2	88	77	HS0002931	AI038455 Homo sapi	208	37	1.2	267	35	AI130479	SWOVL3CAN
136	37	1.2	98	38	AI349279	AI349279 ta75d08.x	209	37	1.2	269	72	AW424164	AW424164 sh61q07.y
137	37	1.2	100	36	AI210082	AI210082 g9c01a1.r	c 210	37	1.2	270	34	AI007951	EST202402
138	37	1.2	102	36	AI676803	AI676803 oa45h12.s	211	37	1.2	270	70	AW311267	sg34f02.y
139	37	1.2	104	69	AW191003	AW191003 xl67a03.x	212	37	1.2	272	48	AU058132	AU058132
140	37	1.2	109	73	AW481860	AW481860 38894 MAR	213	37	1.2	273	69	AW185815	se60a06.y
141	37	1.2	110	43	AI702343	AI702343 tz66b01.x	c 214	37	1.2	274	41	AI596861	vs32a05.x
142	37	1.2	112	38	AI340552	AI340552 tb30f07.x	215	37	1.2	275	37	AI251573	qv45a10.x
143	37	1.2	112	38	AI340553	AI340553 tb30h09.x	c 216	37	1.2	275	41	AI591131	tw23f11.x
144	37	1.2	125	41	AI538850	AI538850 tp74e02.x	217	37	1.2	275	64	AW126272	N100379e
145	37	1.2	127	21	AA098579	AA098579 mm83f05.r	218	37	1.2	275	72	AW459492	AW459492 sh41h11.y
146	37	1.2	128	38	AI335363	AI335363 tb79a05.x	219	37	1.2	276	48	AU085927	AU085927
147	37	1.2	131	38	AI343091	AI343091 tb04b01.x	220	37	1.2	278	30	AA734166	vs19h02.r
148	37	1.2	134	64	AW156654	AW156654 se29a07.y	c 221	37	1.2	278	64	AW129218	xf21e11.x
149	37	1.2	136	46	AI950675	AI950675 wx52h04.x	222	37	1.2	282	72	AW433419	707068G03
150	37	1.2	137	22	AA191932	AA191932 rs04b02.r	c 223	37	1.2	284	41	AI545574	fb70f08.x
151	37	1.2	138	37	AI273966	AI273966 qu90h10.x	224	37	1.2	284	47	AU033071	AU033071
152	37	1.2	145	42	AI644528	AI644528 603005G08	c 225	37	1.2	286	33	AA987387	AA987387 qg86g10.s
153	37	1.2	148	43	AI687016	AI687016 tp81h10.x	c 226	37	1.2	289	46	AI932547	AI932547 wo03f11.x
154	37	1.2	154	73	AW099488	AW099488 sd41d09.y	227	37	1.2	290	81	AW306179	se47a01.y
155	37	1.2	154	73	AW480921	AW480921 33577 MAR	228	37	1.2	290	81	C93414	C93414 C93414 D1ct
156	37	1.2	155	22	AA162707	AA162707 mn42c02.r	c 229	37	1.2	291	31	AA807269	oc33d07.s
157	37	1.2	155	47	AI036886	AI036886 DKF2p564P	c 230	37	1.2	291	40	AI492950	qz47a07.x
158	37	1.2	157	64	AW156429	AW156429 se26a12.y	231	37	1.2	291	46	AI944478	bs01h05.y
159	37	1.2	161	29	AA662912	AA662912 nu92b02.s	c 232	37	1.2	295	31	AA837029	og09e06.s
160	37	1.2	161	45	AI1886062	AI1886062 wn15a04.x	233	37	1.2	296	22	AA185548	mu48h04.r
161	37	1.2	162	42	AI666647	AI666647 mt53b05.x	c 234	37	1.2	297	91	W72952	zd59a09.s1
162	37	1.2	169	32	AA873345	AA873345 ob64g08.s	c 235	37	1.2	298	43	AI739780	606051A01
163	37	1.2	170	31	AA827658	AA827658 od01c02.s	c 236	37	1.2	302	64	AW132999	se13b01.y
164	37	1.2	171	26	AA414389	AA414389 vd07f10.s	237	37	1.2	302	79	AW640213	bl92e11.w
165	37	1.2	172	22	AA162226	AA162226 mn44g01.r	238	37	1.2	302	81	C90617	C90617 D1ct
166	37	1.2	176	63	AW100657	AW100657 sd58c08.y	c 239	37	1.2	313	46	AI917106	ts74d03.x
167	37	1.2	182	28	AA606861	AA606861 vm83b09.r	240	37	1.2	315	27	AA517991	vl18c06.r
168	37	1.2	183	89	TO1212	TO1212 wEST01933 E	c 241	37	1.2	315	46	AI926024	wo41a05.x
169	37	1.2	184	72	AW414544	AW414544 48060 MAR	c 242	37	1.2	322	68	AU075938	AU075938
170	37	1.2	189	72	AW432419	AW432419 sh73g02.y	c 243	37	1.2	322	69	AW193598	AW193598 km18b01.x
171	37	1.2	193	71	AW396099	AW396099 sh02c11.y	244	37	1.2	323	30	AA709810	vt39a06.r
172	37	1.2	194	40	AI492521	AI492521 ti29a02.x	245	37	1.2	326	72	AW458018	sh77b11.y
173	37	1.2	194	63	AW101857	AW101857 sd71e12.y	c 246	37	1.2	327	31	AA775857	ad17f03.s
174	37	1.2	199	20	AA017992	AA017992 mh45a07.r	c 247	37	1.2	327	42	AI649540	603006A06
175	37	1.2	199	47	AU012951	AU012951 EHST101	248	37	1.2	329	20	AA024135	mn97a02.r
176	37	1.2	201	73	AW503171	AW503171 UI-HF-BNO	249	37	1.2	330	91	WB4086	T2987 NVA74
177	37	1.2	204	31	AA836664	AA836664 of53a03.s	250	37	1.2	331	23	AA260469	va94d04.r
178	37	1.2	205	41	AI540361	AI540361 tp63h08.x	c 251	37	1.2	337	28	AA558973	nl10c02.s
179	37	1.2	207	50	AV173975	AV173975 AV173975	252	37	1.2	337	81	C93386	C93386 D1ct
180	37	1.2	211	45	AI862245	AI862245 tb77c12.x	253	37	1.2	337	81	C93399	C93399 D1ct
181	37	1.2	211	62	AW026635	AW026635 ww45e08.x	c 254	37	1.2	338	46	AI927899	wp93b12.x
182	37	1.2	212	38	AI345358	AI345358 tb68e12.x	255	37	1.2	338	71	AW396703	sh29f10.y
183	37	1.2	212	87	N98121	N98121 2213C3 czap	c 256	37	1.2	344	31	AA830625	oc50g09.s
184	37	1.2	213	20	AA018016	AA018016 mh45c07.r	c 257	37	1.2	344	38	AI331810	fb01q12.x
185	37	1.2	217	46	AI946027	AI946027 bs19g10.y	c 258	37	1.2	346	32	AA918425	ol70h02.s
186	37	1.2	221	27	AA523466	AA523466 ni63h06.s	c 259	37	1.2	346	63	AW100829	ad17h02.s
187	37	1.2	225	64	AW132200	AW132200 sd67e05.y	c 260	37	1.2	353	30	AA767170	oc88h08.s
188	37	1.2	233	63	AW080611	AW080611 xc44a06.x	261	37	1.2	356	80	AW733209	sk71d05.x
189	37	1.2	236	23	AA274698	AA274698 vc05b05.r	c 262	37	1.2	357	35	AI086823	oz57g01.x
190	37	1.2	238	73	AA290000	AA290000 vb50b07.r	263	37	1.2	357	72	AW415996	AW415996 50650 MAR
191	37	1.2	239	73	AW483334	AW483334 52580 MAR	264	37	1.2	359	81	C91502	C91502 D1ct
192	37	1.2	240	48	AU072476	AU072476 AU072476	265	37	1.2	361	92	X73770	CHESTW067 G
193	37	1.2	240	48	AU074000	AU074000 AU074000	266	37	1.2	362	47	AU071104	AU071104
194	37	1.2	244	42	AI650711	AI650711 wb25g07.x	267	37	1.2	363	47	AU029668	AU029668
195	37	1.2	245	63	AW100838	AW100838 sd62a06.y	c 268	37	1.2	363	70	AW288760	707009D01
196	37	1.2	245	67	N69347	N69347 yz43e08.s1	c 269	37	1.2	364	35	AI073388	oo13a08.x
197	37	1.2	248	64	AW119367	AW119367 sd45c07.y	c 270	37	1.2	366	38	AI341886	AI341886 qs98b09.x
198	37	1.2	251	88	R65559	R65559 14063 Lambd	271	37	1.2	366	92	Z98088	Z98088
199	37	1.2	253	40	AI495706	AI495706 sb15c08.y	c 272	37	1.2	369	39	AI395143	AI395143
200	37	1.2	253	64	AW119354	AW119354 sd45a12.y	273	37	1.2	369	70	AW288761	707009D01
201	37	1.2	255	63	AW065230	AW065230 614038B03	274	37	1.2	369	81	C99552	C99552 Rice

275	37	1.2	370	47	AU031086	AU031086	AU031086	C 348	37	1.2	501	63	AW058793	AW058793	fe47h06.x
276	37	1.2	371	30	AA760782	nz14c04.s	AA760782	C 349	37	1.2	502	69	AW180895	MgA0006r	
277	37	1.2	372	30	AI271716	qj87g06.x	AI271716	C 350	37	1.2	511	29	AA633412	np78e10.s	
278	37	1.2	375	47	AU029637	AU029637	AU029637	C 351	37	1.2	512	38	AI382230	td04a05.x	
279	37	1.2	378	31	AA804629	ob98c12.s	AA804629	C 352	37	1.2	513	31	AA829330	od05h09.s	
280	37	1.2	378	64	AW129589	AW129589	AW129589	C 353	37	1.2	513	31	AA830374	oc79h02.s	
281	37	1.2	381	23	AA259538	va49h11.r	AA259538	C 354	37	1.2	516	43	AI738199	oc64g01.s	
282	37	1.2	383	30	AA709128	zf92h07.s	AA709128	C 355	37	1.2	518	45	AI890391	vm84g04.x	
283	37	1.2	383	39	AI395176	MA002571.	AI395176	C 356	37	1.2	521	30	AA766930	oc88f12.s	
284	37	1.2	384	48	AU057780	AU057780	AU057780	C 357	37	1.2	532	48	AU082640	AU082640	
285	37	1.2	385	23	AA243164	zr26d01.s	AA243164	C 358	37	1.2	533	48	AU082139	AU082139	
286	37	1.2	385	79	AW669208	112464.MA	AW669208	C 359	37	1.2	537	31	AA835963	oc79h02.s	
287	37	1.2	387	48	AU071229	AU071229	AU071229	C 360	37	1.2	538	21	AA081411	zn34h10.s	
288	37	1.2	389	39	AI393637	tg59e12.x	AI393637	C 361	37	1.2	539	42	AI666130	606005C10	
289	37	1.2	389	79	AW633117	AW633117	AW633117	C 362	37	1.2	547	28	AA547940	nk51h07.s	
290	37	1.2	397	63	AW078751	AW078751	AW078751	C 363	37	1.2	547	28	AA594974	no37b09.s	
291	37	1.2	399	31	AA777002	AI226246	AA777002	C 364	37	1.2	549	43	AI728861	BNLGH1118	
292	37	1.2	399	37	AI226246	ue88h04.y	AI226246	C 365	37	1.2	552	48	AU037695	AU037695	
293	37	1.2	399	74	AW568851	sl73a10.y	AW568851	C 366	37	1.2	553	47	AU034075	AU034075	
294	37	1.2	400	31	AA824581	AW824581	AA824581	C 367	37	1.2	563	32	AA898161	NCM4H9T7	
295	37	1.2	400	35	AI131746	AI131746	AI131746	C 368	37	1.2	571	63	AW067355	683022G06	
296	37	1.2	401	80	AW733536	AW733536	AW733536	C 369	37	1.2	574	47	AU032687	AU032687	
297	37	1.2	404	42	AI624315	ts67c05.x	AI624315	C 370	37	1.2	574	48	AU083377	AU083377	
298	37	1.2	404	80	AW706146	AW706146	AW706146	C 371	37	1.2	575	39	AI401302	AI401302	
299	37	1.2	406	23	AA233530	zr30g06.s	AA233530	C 372	37	1.2	577	27	AA522882	AI41B03.S	
300	37	1.2	409	41	AI579868	1028.NC-1	AI579868	C 373	37	1.2	580	119	AZ066515	RPCL-23-4	
301	37	1.2	409	45	AI888650	wn22d07.x	AI888650	C 374	37	1.2	582	46	AI958190	cg91h10.y	
302	37	1.2	410	31	AA805093	AA805093	AA805093	C 375	37	1.2	583	32	AA901876	NCC4C8T7	
303	37	1.2	410	35	AI090233	AI090233	AI090233	C 376	37	1.2	588	43	AI730627	BNLGH1742	
304	37	1.2	411	44	AI762931	wh94e08.x	AI762931	C 377	37	1.2	597	30	AA701808	PMIFG_136	
305	37	1.2	417	79	AW633129	AI64b02.x	AW633129	C 378	37	1.2	603	47	AU032494	AU032494	
306	37	1.2	418	37	AI255423	u194d02.x	AI255423	C 379	37	1.2	604	36	AI207574	HA2938.Hu	
307	37	1.2	418	37	AI303093	AI303093	AI303093	C 380	37	1.2	609	42	AI680453	tw81e01.x	
308	37	1.2	422	27	AA489687	AA489687	AA489687	C 381	37	1.2	610	48	AU038887	AU038887	
309	37	1.2	425	31	AA836303	od36d04.s	AA836303	C 382	37	1.2	611	36	AI172702	EST179.Ma	
310	37	1.2	429	28	AA569361	th89f12.s	AA569361	C 383	37	1.2	611	36	AI174996	HA2736.Hu	
311	37	1.2	430	30	AA703338	zj11d06.s	AA703338	C 384	37	1.2	612	32	AA873373	Oh69b08.S	
312	37	1.2	430	81	C99609	C99609	C99609	C 385	37	1.2	614	64	AI108250	um20a02.x	
313	37	1.2	433	31	AA830841	AI830841	AA830841	C 386	37	1.2	614	64	AI816377	au47c02.x	
314	37	1.2	436	62	AW019898	df01c10.y	AW019898	C 387	37	1.2	617	63	AW088253	xc99f01.x	
315	37	1.2	439	47	AL042490	DKFzp434G	AL042490	C 388	37	1.2	627	48	AU068669	AU068669	
316	37	1.2	448	36	AI160859	qb65f08.x	AI160859	C 389	37	1.2	632	27	AA504326	aa61e05.s	
317	37	1.2	448	39	AI431915	tl26f10.x	AI431915	C 390	37	1.2	634	81	C96971	C96971	
318	37	1.2	448	69	AW188426	xj98f04.x	AW188426	C 391	37	1.2	646	64	AW164210	se23c10.y	
319	37	1.2	451	31	AA813963	AI813963	AA813963	C 392	37	1.2	647	47	AU030767	AU030767	
320	37	1.2	451	47	AU029647	AU029647	AU029647	C 393	37	1.2	656	70	AW265928	L30-2458T	
321	37	1.2	452	31	AA815048	AA815048	AA815048	C 394	37	1.2	669	47	AL041154	DRFzp434F	
322	37	1.2	456	81	C91915	C91915	C91915	C 395	37	1.2	671	42	AI624433	ts29g03.x	
323	37	1.2	457	63	AW080747	AW080747	AW080747	C 396	37	1.2	676	80	AW725836	GA_Ea001	
324	37	1.2	458	48	AU057569	AU057569	AU057569	C 397	37	1.2	679	48	AU056612	AU056612	
325	37	1.2	458	79	AW632966	bl02c09.x	AW632966	C 398	37	1.2	684	39	AI395775	MA005127.	
326	37	1.2	465	39	AI440525	tc89g12.x	AI440525	C 399	37	1.2	712	48	AU075852	AU075852	
327	37	1.2	465	46	AI974961	EST269555	AI974961	C 400	37	1.2	718	40	AI478986	tn29d01.x	
328	37	1.2	467	48	AU057247	AU057247	AU057247	C 401	37	1.2	726	27	AA528799	nf52b07.s	
329	37	1.2	470	28	AA602351	np47g01.s	AA602351	C 402	37	1.2	726	81	C99640	C99640	
330	37	1.2	470	47	AU033739	AU033739	AU033739	C 403	37	1.2	729	62	AW019812	fg46h04.x	
331	37	1.2	474	81	C98042	C98042	C98042	C 404	37	1.2	730	47	AL036742	DKFzp564I	
332	37	1.2	476	39	AI395813	MA005189.	AI395813	C 405	37	1.2	733	62	AV405504	AV405504	
333	37	1.2	476	44	AI829893	wj47a10.x	AI829893	C 406	37	1.2	991	122	CMS00JW8	AL077118.Drosophi	
334	37	1.2	483	23	AA277199	vb91g11.r	AA277199	C 407	36	1.2	49	27	AA526728	n191d10.s	
335	37	1.2	486	37	AI297124	lp11346.s	AI297124	C 408	36	1.2	49	28	AA608271	vn61f09.r	
336	37	1.2	486	47	AU029499	AU029499	AU029499	C 409	36	1.2	54	28	AA601314	tt15f06.s	
337	37	1.2	486	63	AW041017	EST283881	AW041017	C 410	36	1.2	54	42	AI633751	tt128b05.x	
338	37	1.2	487	47	AL045439	DKFzp434F	AL045439	C 411	36	1.2	60	72	AW409793	fh02d08.y	
339	37	1.2	489	79	AW642784	cm22e06.w	AW642784	C 412	36	1.2	64	80	AW715605	g95g02nm.I	
340	37	1.2	489	78	AA624284	AA624284	AA624284	C 413	36	1.2	69	29	AA639635	nq83e02.s	
341	37	1.2	494	32	AA898370	NCC2E1T7	AA898370	C 414	36	1.2	72	43	AI687854	tp89a03.x	
342	37	1.2	494	46	AI961384	wt17h04.x	AI961384	C 415	36	1.2	72	31	AA826924	ob52b03.s	
343	37	1.2	495	31	AA834558	od63d09.s	AA834558	C 416	36	1.2	79	89	T25558	EST00591.Un	
344	37	1.2	496	23	AA274427	pb07f09.r	AA274427	C 417	36	1.2	80	44	AI810694	tl19h09.x	
345	37	1.2	500	41	AI596140	uk25a02.x	AI596140	C 418	36	1.2	84	38	AI335218	ta93b09.x	
346	37	1.2	500	48	AU037533	AU037533	AU037533	C 419	36	1.2	84	84	AU053727	AU053727	
347	37	1.2	501	44	AI819127	wj79d02.x	AI819127	C 420	36	1.2	84	89	T25553	EST00586.Un	

c 421	36	1.2	85	28	AA606915	vm91c01.r	c 494	36	1.2	141	30	AA720874	AA720874
c 422	36	1.2	86	35	AI086378	oz44c01.x	c 495	36	1.2	142	39	AI420269	AI420269
c 423	36	1.2	87	77	HSM001845	AI037518 Homo sapi	496	36	1.2	142	62	AW022357	AW022357
c 424	36	1.2	88	28	AA590663	vm24h10.i	497	36	1.2	142	72	AW423982	AW423982
c 425	36	1.2	88	64	AW151979	xf71c01.x	c 498	36	1.2	143	21	AA106528	AA106528
c 426	36	1.2	89	41	AI559863	co32e12.x	c 499	36	1.2	143	27	AA106528	AA106528
c 427	36	1.2	90	30	AA708028	AW708028 zg05e01.s	500	36	1.2	143	21	AA106528	AA106528
c 428	36	1.2	90	91	W48671	W48671 zg45d06.s1	c 501	36	1.2	144	28	AA580663	AA580663
c 429	36	1.2	91	41	AI543508	AI543508 tj23c08.x	c 502	36	1.2	144	31	AA814782	AA814782
c 430	36	1.2	91	41	AI590755	AI590755 tw18d10.x	c 503	36	1.2	146	39	AI437025	AI437025
c 431	36	1.2	93	46	AI960759	AI960759 sc90a09.y	c 504	36	1.2	146	39	AI444985	AI444985
c 432	36	1.2	95	79	AW683162	AW683162 NF008C12L	c 505	36	1.2	146	39	AI444985	AI444985
c 433	36	1.2	97	38	AI321720	AI321720 e2d07nm.f	c 506	36	1.2	147	79	AW632958	AW632958
c 434	36	1.2	99	20	AA026667	AA026667 te93h01.s	c 507	36	1.2	147	79	AW632958	AW632958
c 435	36	1.2	99	43	AI689157	AI689157 tx92f10.x	c 508	36	1.2	149	34	AI1017427	AI1017427
c 436	36	1.2	100	28	AA590470	AA590470 vm24c02.i	c 509	36	1.2	149	38	AI1345608	AI1345608
c 437	36	1.2	100	28	AA606796	AA606796 vm90f07.i	c 510	36	1.2	149	62	AW020763	AW020763
c 438	36	1.2	100	72	AW438968	AW438968 77621.MAR	c 511	36	1.2	149	72	AW417850	AW417850
c 439	36	1.2	101	28	AA555145	AA555145 nl07g04.s	c 512	36	1.2	150	42	AI652921	AI652921
c 440	36	1.2	101	39	AI1415241	AI1415241 mdl4d03.x	c 513	36	1.2	150	47	AL039118	AL039118
c 441	36	1.2	102	79	AW681730	AW681730 EST00448	c 514	36	1.2	150	63	AW101452	AW101452
c 442	36	1.2	102	80	C21335	C21335 HUMGS000333	c 515	36	1.2	151	42	AI620145	AI620145
c 443	36	1.2	103	47	AL048499	AL048499 DK52p586M	c 516	36	1.2	151	44	AI801558	AI801558
c 444	36	1.2	104	48	AW053720	AW053720 AU053720	c 517	36	1.2	151	64	AW163834	AW163834
c 445	36	1.2	105	22	AA216932	AA216932 mv75c12.i	c 518	36	1.2	152	28	AA582366	AA582366
c 446	36	1.2	105	86	H89138	H89138 yw27a08.i1	c 519	36	1.2	152	72	AW424113	AW424113
c 447	36	1.2	106	22	AA172938	AA172938 mt01g10.i	c 520	36	1.2	153	81	AI590235	AI590235
c 448	36	1.2	106	72	AW416451	AW416451 51810.MAR	c 521	36	1.2	153	41	C90066	C90066
c 449	36	1.2	108	48	AV021502	AV021502 AV021502	c 522	36	1.2	154	21	AA145131	AA145131
c 450	36	1.2	109	42	AI659651	AI659651 tu25b02.x	c 523	36	1.2	154	70	AW315888	AW315888
c 451	36	1.2	111	30	AA761557	AA761557 n223c03.s	c 524	36	1.2	154	79	AW633164	AW633164
c 452	36	1.2	111	37	AI289394	AI289394 qw21h08.x	c 525	36	1.2	155	23	AA242068	AA242068
c 453	36	1.2	111	38	AI345471	AI345471 tb82f09.x	c 526	36	1.2	155	37	AI280010	AI280010
c 454	36	1.2	112	42	AI612446	AI612446 486085C04	c 527	36	1.2	155	46	AI938906	AI938906
c 455	36	1.2	112	70	AW307653	AW307653 1387.MAR	c 528	36	1.2	155	73	AW479501	AW479501
c 456	36	1.2	113	64	AW161202	AW161202 au79g05.y	c 529	36	1.2	156	41	AI561356	AI561356
c 457	36	1.2	113	72	AW430509	AW430509 70135.MAR	c 530	36	1.2	156	48	AW307419	AW307419
c 458	36	1.2	113	77	HSM002662	HSM002662 Homo sapi	c 531	36	1.2	156	69	AW213895	AW213895
c 459	36	1.2	114	39	AI433008	AI433008 th42e09.x	c 532	36	1.2	156	73	AW479507	AW479507
c 460	36	1.2	117	37	AI289791	AI289791 qw12c03.x	c 533	36	1.2	156	73	AW479507	AW479507
c 461	36	1.2	117	73	AW486832	AW486832 77218.MAR	c 534	36	1.2	157	45	AI686163	AI686163
c 462	36	1.2	118	33	AA939202	AA939202 oi99h05.s	c 535	36	1.2	157	64	AW164657	AW164657
c 463	36	1.2	118	39	AI444981	AI444981 ti93d01.x	c 536	36	1.2	157	73	AW485902	AW485902
c 464	36	1.2	119	63	AW089152	AW089152 xc79g07.x	c 537	36	1.2	157	80	C23819	C23819
c 465	36	1.2	119	72	AW426534	AW426534 60821.MAR	c 538	36	1.2	158	37	AI289386	AI289386
c 466	36	1.2	120	70	AW264380	AW264380 NF032G12S	c 539	36	1.2	159	31	AA811586	AA811586
c 467	36	1.2	121	31	AA830333	AA830333 oc50b09.s	c 540	36	1.2	159	39	AI394443	AI394443
c 468	36	1.2	121	63	AW065627	AW065627 614056D11	c 541	36	1.2	159	42	AI559504	AI559504
c 469	36	1.2	121	72	AW436766	AW436766 77229.MAR	c 542	36	1.2	159	64	AW151136	AW151136
c 470	36	1.2	122	22	AA149757	AA149757 zo01f01.s	c 543	36	1.2	159	71	AW336889	AW336889
c 471	36	1.2	123	79	AW681481	AW681481 EST000159	c 544	36	1.2	159	81	C99492	C99492
c 472	36	1.2	125	38	AI361271	AI361271 qy42g06.x	c 545	36	1.2	160	22	AA151875	AA151875
c 473	36	1.2	125	79	AW642236	AW642236 cm17d03.w	c 546	36	1.2	160	71	AW327210	AW327210
c 474	36	1.2	127	63	AA028993	AA028993 wv98b09.x	c 547	36	1.2	160	71	AW336921	AW336921
c 475	36	1.2	128	23	AA222777	AA222777 mv99h10.i	c 548	36	1.2	160	89	T04635	T04635
c 476	36	1.2	130	22	AA217439	AA217439 mu94a09.i	c 549	36	1.2	161	37	AI282595	AI282595
c 477	36	1.2	130	42	AI611647	AI611647 tt37g05.x	c 550	36	1.2	161	64	AW118252	AW118252
c 478	36	1.2	130	48	AW037699	AW037699 AU037699	c 551	36	1.2	161	72	AW416297	AW416297
c 479	36	1.2	131	33	AA933767	AA933767 om59h09.s	c 552	36	1.2	162	39	AI446178	AI446178
c 480	36	1.2	131	48	AW037779	AW037779 AU037779	c 553	36	1.2	162	47	AW032238	AW032238
c 481	36	1.2	131	77	HSM006247	HSM006247 Homo sapi	c 554	36	1.2	162	64	AW132939	AW132939
c 482	36	1.2	132	43	AI702797	AI702797 we07e02.x	c 555	36	1.2	163	44	AI802244	AI802244
c 483	36	1.2	135	30	AA737267	AA737267 o2e05.s	c 556	36	1.2	164	42	AI613028	AI613028
c 484	36	1.2	135	41	AI560511	AI560511 tq46g06.x	c 557	36	1.2	164	43	AI758547	AI758547
c 485	36	1.2	135	42	AI632096	AI632096 ts79g12.x	c 558	36	1.2	164	63	AW079012	AW079012
c 486	36	1.2	136	37	AI271763	AI271763 qj77d05.x	c 559	36	1.2	164	81	C92873	C92873
c 487	36	1.2	136	72	AW429931	AW429931 68536.MAR	c 560	36	1.2	165	30	AA766539	AA766539
c 488	36	1.2	136	72	AW432459	AW432459 sh74c12.y	c 561	36	1.2	165	40	AI471325	AI471325
c 489	36	1.2	137	72	AA429406	AA429406 67718.MAR	c 562	36	1.2	166	71	AW396791	AW396791
c 490	36	1.2	140	32	AA878532	AA878532 oj19d09.s	c 563	36	1.2	166	73	AW485427	AW485427
c 491	36	1.2	140	38	AI364783	AI364783 qul18e12.x	c 564	36	1.2	166	74	AW570333	AW570333
c 492	36	1.2	140	48	AU056957	AU056957 AU056957	c 565	36	1.2	167	20	AA024151	AA024151
c 493	36	1.2	140	48	AU056957	AU056957 AU056957	c 566	36	1.2	167	30	AA749425	AA749425

567	36	1.2	167	47	AL079802	AL079802	DKE2p4340	c	640	190	64	AW167460	
568	36	1.2	168	22	AA189350	AA189350	mt85c01.r	641	36	1.2	190	72	AW430315
569	36	1.2	168	23	AA276966	AA276966	vc42h11.r	642	36	1.2	190	74	AW573456
c 570	36	1.2	168	45	AI866820	AI866820	tz53f09.x	c 643	36	1.2	191	37	AI244105
c 571	36	1.2	168	72	AW428528	AW428528	667111.MAR	c 644	36	1.2	191	47	AL041772
c 572	36	1.2	169	36	AI167912	AI167912	ok29b07.x	645	36	1.2	191	74	AW569034
c 573	36	1.2	169	38	AI345010	AI345010	ta95b02.x	c 646	36	1.2	192	32	AA916898
c 574	36	1.2	169	41	AI545595	AI545595	fb67b10.x	c 647	36	1.2	192	32	AI275178
575	36	1.2	169	47	AI119791	AI119791	DKE2p761F	c 648	36	1.2	192	37	AI282031
576	36	1.2	170	21	AA116459	AA116459	mp95d12.r	c 649	36	1.2	192	42	AI622052
c 577	36	1.2	171	47	AL040077	AL040077	DKE2p434A	c 650	36	1.2	192	45	AI856699
578	36	1.2	172	29	AA638609	AA638609	vo55c04.r	651	36	1.2	193	23	AA266283
c 579	36	1.2	172	42	AI651093	AI651093	wa97b11.x	c 652	36	1.2	193	27	AA490318
580	36	1.2	172	63	AW100667	AW100667	fs58d11.y	c 653	36	1.2	193	30	AA720850
581	36	1.2	172	71	AW346438	AW346438	28052.MAR	c 654	36	1.2	193	31	AA811663
582	36	1.2	173	35	AI105939	AI105939	cb03h10.t	c 655	36	1.2	193	35	AI144106
c 583	36	1.2	173	41	AI583023	AI583023	tr99d10.x	c 656	36	1.2	193	37	AI280528
584	36	1.2	173	47	AL120706	AL120706	DKE2p762G	c 657	36	1.2	193	64	AW168700
585	36	1.2	174	64	AW164430	AW164430	se72d09.y	c 658	36	1.2	193	117	AA922522
c 586	36	1.2	174	31	AA836253	AA836253	od17c12.s	c 659	36	1.2	194	22	AA147918
c 587	36	1.2	174	41	AI606084	AI606084	ve97g04.x	c 660	36	1.2	194	29	AA693354
c 588	36	1.2	174	62	AW020529	AW020529	df11b04.y	c 661	36	1.2	194	37	AI285817
c 589	36	1.2	174	63	AW082510	AW082510	xb52a09.x	c 662	36	1.2	194	40	AI471278
590	36	1.2	174	72	AW417645	AW417645	54545.MAR	c 663	36	1.2	195	36	AI174394
591	36	1.2	174	72	AW430301	AW430301	69657.MAR	c 664	36	1.2	195	41	AI540416
c 592	36	1.2	176	30	AA749184	AA749184	nylle06.s	c 665	36	1.2	195	43	AI689340
c 593	36	1.2	176	42	AI661526	AI661526	va49g02.x	c 666	36	1.2	195	44	AI824576
c 594	36	1.2	176	46	AI952080	AI952080	wx46g01.x	c 667	36	1.2	196	22	AA173255
595	36	1.2	176	63	AW102540	AW102540	sd60e12.y	c 668	36	1.2	196	35	AI105952
c 596	36	1.2	177	30	AA764903	AA764903	nz76h11.s	669	36	1.2	196	81	F00107
c 597	36	1.2	177	30	AA768725	AA768725	oc86g07.s	670	36	1.2	196	89	T40336
c 598	36	1.2	177	46	AI934065	AI934065	wg96b01.x	671	36	1.2	197	38	AI376494
599	36	1.2	177	71	AA396827	AA396827	sg53g02.y	c 672	36	1.2	197	45	AI889872
600	36	1.2	177	72	AA414476	AA414476	48167.MAR	673	36	1.2	198	22	AA189279
601	36	1.2	178	25	AA389924	AA389924	vb28g06.r	674	36	1.2	198	47	AL134242
c 602	36	1.2	178	31	AA806857	AA806857	OC29e05.s	675	36	1.2	198	63	AAW101681
603	36	1.2	178	47	AJ273583	AJ273583	OC29e05.s	c 676	36	1.2	199	30	AA713657
604	36	1.2	179	20	AA009346	AA009346	TgESTzz08	c 677	36	1.2	199	44	AI796143
c 605	36	1.2	179	30	AA767252	AA767252	oc89f08.s	c 678	36	1.2	199	63	AW087938
c 606	36	1.2	180	37	AI241824	AI241824	qu72d09.x	679	36	1.2	199	79	AA654681
c 607	36	1.2	180	63	AW089293	AW089293	xd3e07.x	c 680	36	1.2	200	31	AA837290
c 608	36	1.2	181	39	AI401088	AI401088	th22c02.x	c 681	36	1.2	200	32	AA911452
c 609	36	1.2	181	63	AI701868	AI701868	tl91a03.x	c 682	36	1.2	201	36	AI223980
c 610	36	1.2	181	69	AW227579	AW227579	up11b08.x	c 683	36	1.2	201	37	AI284131
611	36	1.2	181	73	AA477996	AA477996	18034.MAR	c 684	36	1.2	201	45	AI890451
c 612	36	1.2	182	45	AI886201	AI886201	wn17h08.x	c 685	36	1.2	201	63	AW103317
c 613	36	1.2	182	47	AL119863	AL119863	DKE2p761L	c 686	36	1.2	201	64	AW168341
614	36	1.2	182	64	AW156216	AW156216	se21a09.y	687	36	1.2	201	72	AA423389
c 615	36	1.2	182	73	AA479525	AA479525	25756.MAR	688	36	1.2	202	22	AA198860
c 616	36	1.2	183	22	AA179931	AA179931	zp56h09.s	689	36	1.2	203	23	AA274632
c 617	36	1.2	183	22	AA265647	AA265647	mw14c11.r	690	36	1.2	203	23	AA275299
c 618	36	1.2	183	32	AA860080	AA860080	ak45c09.s	691	36	1.2	203	23	AA275355
c 619	36	1.2	183	63	AW029009	AW029009	wv98d11.x	692	36	1.2	203	63	AW100931
c 620	36	1.2	183	72	AA447056	AA447056	88151.MAR	693	36	1.2	204	23	AA289196
c 621	36	1.2	184	91	W32269	W32269	zc04h06.sl	c 694	36	1.2	204	40	AI493133
c 622	36	1.2	184	29	AA632400	AA632400	np67f09.s	c 695	36	1.2	204	44	AI768459
623	36	1.2	184	30	AA710574	AA710574	vt43c11.r	c 696	36	1.2	204	63	AW055261
c 624	36	1.2	184	31	AA805708	AA805708	ns42c02.s	c 697	36	1.2	205	23	AA251340
625	36	1.2	184	63	AW102497	AW102497	sd60a02.y	698	36	1.2	205	81	C90715
626	36	1.2	185	37	AI252901	AI252901	qv31d08.x	699	36	1.2	206	21	AA125000
627	36	1.2	185	71	AA336628	AA336628	21730.MAR	700	36	1.2	206	48	AU038875
c 628	36	1.2	186	23	AA258156	AA258156	zs35a12.s	701	36	1.2	206	73	AA483853
c 629	36	1.2	186	80	AW718326	AW718326	14g11nm.f	702	36	1.2	207	31	AA795017
630	36	1.2	187	21	AA145886	AA145886	mr33b12.r	c 703	36	1.2	207	37	AI277260
631	36	1.2	187	28	AA611254	AA611254	vo51a12.r	704	36	1.2	207	37	AI284542
c 632	36	1.2	187	42	AI620042	AI620042	ty47f08.x	c 705	36	1.2	207	38	AI352497
633	36	1.2	187	47	AL044623	AL044623	DKE2p4340	706	36	1.2	207	63	AW101865
634	36	1.2	187	63	AW099176	AW099176	sd35f12.y	707	36	1.2	207	80	AW734306
c 635	36	1.2	187	64	AW129264	AW129264	xf22f05.x	708	36	1.2	208	27	AA512135
636	36	1.2	188	70	AW307553	AW307553	sf58d04.y	709	36	1.2	208	33	AF074644
637	36	1.2	188	23	AA274350	AA274350	TgESTzz25	c 710	36	1.2	208	34	AI024643
638	36	1.2	189	47	AL036214	AL036214	DKE2p564H	711	36	1.2	208	35	AI110737
c 639	36	1.2	190	33	AA959780	AA959780	vw55b09.s	c 712	36	1.2	208	37	AI266652

c 713	c 713	36	1.2	208	40	AI522772	AI522772	fb62c01.x	c 786	36	1.2	222	28	AA583860	nn60cl2.x
c 714	c 714	36	1.2	208	42	AI640130	AI640130	wa29e02.x	c 787	36	1.2	222	39	AI440430	AI440430
c 715	c 715	36	1.2	209	41	AI571909	AI571909	tr71b02.x	c 788	36	1.2	222	34	AW132524	AW132524
c 716	c 716	36	1.2	209	43	AI683776	AI683776	tw53h07.x	c 789	36	1.2	222	37	AW425645	AW425645
c 717	c 717	36	1.2	209	48	AU074565	AU074565	AU074565	c 790	36	1.2	223	72	AI250342	AI250342
c 718	c 718	36	1.2	209	63	AW102550	AW102550	sd60f12.y	c 791	36	1.2	223	37	AI250342	AI250342
c 719	c 719	36	1.2	209	69	AW193280	AW193280	x172h08.x	c 792	36	1.2	224	33	AA936133	AA936133
c 720	c 720	36	1.2	209	71	AW345987	AW345987	26929.MAR	c 793	36	1.2	224	41	AI559551	AI559551
c 721	c 721	36	1.2	209	71	AW396627	AW396627	sh26a01.y	c 794	36	1.2	224	71	AW353417	AW353417
c 722	c 722	36	1.2	209	72	AW426942	AW426942	61884.MAR	c 795	36	1.2	224	72	AW415603	AW415603
c 723	c 723	36	1.2	209	72	AW429927	AW429927	68526.MAR	c 796	36	1.2	225	30	AA714186	AA714186
c 724	c 724	36	1.2	210	23	AA235975	AA235975	z505e01.s	c 797	36	1.2	225	37	AI289503	AI289503
c 725	c 725	36	1.2	210	28	AA587817	AA587817	nj08h02.s	c 798	36	1.2	225	47	AL040459	AL040459
c 726	c 726	36	1.2	210	71	AW396612	AW396612	255a04.y	c 799	36	1.2	225	47	AL080011	AL080011
c 727	c 727	36	1.2	211	119	AZ078723	AZ078723	RPCI-23-4	c 800	36	1.2	225	64	AW168497	AW168497
c 728	c 728	36	1.2	212	26	AA411768	AA411768	zu02c04.s	c 801	36	1.2	225	70	AW307246	AW307246
c 729	c 729	36	1.2	212	27	AA489198	AA489198	aa57b01.s	c 802	36	1.2	225	72	AW429704	AW429704
c 730	c 730	36	1.2	212	40	AI491817	AI491817	tn95a06.x	c 803	36	1.2	226	22	AA174678	AA174678
c 731	c 731	36	1.2	212	43	AI689557	AI689557	tx84c07.x	c 804	36	1.2	226	40	AI499446	AI499446
c 732	c 732	36	1.2	212	63	AW101775	AW101775	sd70b04.y	c 805	36	1.2	226	63	AW089171	AW089171
c 733	c 733	36	1.2	212	64	AW168563	AW168563	x189b03.x	c 806	36	1.2	226	63	AW090446	AW090446
c 734	c 734	36	1.2	213	20	AA038775	AA038775	mi95f04.r	c 807	36	1.2	226	64	AW151664	AW151664
c 735	c 735	36	1.2	213	32	AA878812	AA878812	cf83b03.s	c 808	36	1.2	226	71	AW353065	AW353065
c 736	c 736	36	1.2	213	34	AI025427	AI025427	qw27g06.s	c 809	36	1.2	226	72	AW458709	AW458709
c 737	c 737	36	1.2	213	37	AI245325	AI245325	qk28e01.x	c 810	36	1.2	226	73	AW485839	AW485839
c 738	c 738	36	1.2	213	45	AI987741	AI987741	vm37g08.x	c 811	36	1.2	227	33	AA937752	AA937752
c 739	c 739	36	1.2	213	69	AW195444	AW195444	xn38d08.x	c 812	36	1.2	227	42	AI648663	AI648663
c 740	c 740	36	1.2	214	26	AA450649	AA450649	vf80c08.r	c 813	36	1.2	227	64	AW170775	AW170775
c 741	c 741	36	1.2	214	43	AI736553	AI736553	sb30c09.y	c 814	36	1.2	228	31	AA833731	AA833731
c 742	c 742	36	1.2	214	48	AU053098	AU053098	AU053098	c 815	36	1.2	228	38	AI354995	AI354995
c 743	c 743	36	1.2	214	64	AW153129	AW153129	se36c12.y	c 816	36	1.2	228	70	AW263347	AW263347
c 744	c 744	36	1.2	215	45	AI874303	AI874303	wm50h10.x	c 817	36	1.2	228	103	AAQ487403	AAQ487403
c 745	c 745	36	1.2	216	20	AA047715	AA047715	zfl15a01.r	c 818	36	1.2	229	29	AA633592	AA633592
c 746	c 746	36	1.2	216	42	AI628651	AI628651	ty77e12.x	c 819	36	1.2	229	30	AA711151	AA711151
c 747	c 747	36	1.2	216	70	AW314680	AW314680	11001.MAR	c 820	36	1.2	229	41	AI583152	AI583152
c 748	c 748	36	1.2	216	71	AW344855	AW344855	26750.MAR	c 821	36	1.2	229	44	AI810736	AI810736
c 749	c 749	36	1.2	216	72	AW429999	AW429999	68710.MAR	c 822	36	1.2	229	72	AW459376	AW459376
c 750	c 750	36	1.2	217	28	AA596329	AA596329	vo30h01.r	c 823	36	1.2	230	20	AA020512	AA020512
c 751	c 751	36	1.2	217	32	AA873481	AA873481	ch77g03.s	c 824	36	1.2	230	43	AI736133	AI736133
c 752	c 752	36	1.2	217	35	AI272994	AI272994	qv63e01.x	c 825	36	1.2	230	63	AW101750	AW101750
c 753	c 753	36	1.2	217	45	AI860051	AI860051	tx72h04.x	c 826	36	1.2	230	63	AW101750	AW101750
c 754	c 754	36	1.2	217	48	AU078401	AU078401	AU078401	c 827	36	1.2	231	23	AA254732	AA254732
c 755	c 755	36	1.2	217	79	AW631928	AW631928	91315.MAR	c 828	36	1.2	231	23	AI252030	AI252030
c 756	c 756	36	1.2	218	30	AA713785	AA713785	nv70b07.s	c 829	36	1.2	231	42	AI636811	AI636811
c 757	c 757	36	1.2	218	30	AA7766368	AA7766368	ca29b05.s	c 830	36	1.2	231	47	AT000896	AT000896
c 758	c 758	36	1.2	218	38	AI3511164	AI3511164	qt11h12.x	c 831	36	1.2	231	90	W14959	W14959
c 759	c 759	36	1.2	218	47	AL045986	AL045986	DKF7p4340	c 832	36	1.2	232	30	AA743445	AA743445
c 760	c 760	36	1.2	218	62	AU007733	AU007733	wc68e01.x	c 833	36	1.2	232	40	AU0073164	AU0073164
c 761	c 761	36	1.2	218	62	AW007938	AW007938	wv47a07.x	c 834	36	1.2	232	70	AW306696	AW306696
c 762	c 762	36	1.2	218	62	AW099572	AW099572	sd43f11.y	c 835	36	1.2	232	70	AW306696	AW306696
c 763	c 763	36	1.2	218	72	AW459746	AW459746	sh91e09.y	c 836	36	1.2	233	45	AI887752	AI887752
c 764	c 764	36	1.2	219	20	AA027470	AA027470	mi05d01.r	c 837	36	1.2	233	63	AW057680	AW057680
c 765	c 765	36	1.2	219	30	AA0709719	AA0709719	vt38a10.r	c 838	36	1.2	233	88	R65859	R65859
c 766	c 766	36	1.2	219	39	AI452993	AI452993	tj46c09.x	c 839	36	1.2	234	33	AA231147	AA231147
c 767	c 767	36	1.2	219	45	AI865070	AI865070	wk07b02.x	c 840	36	1.2	234	47	AL043073	AL043073
c 768	c 768	36	1.2	219	64	AW156260	AW156260	se21f03.y	c 841	36	1.2	234	63	AW100583	AW100583
c 769	c 769	36	1.2	219	64	AW156522	AW156522	se28c01.y	c 842	36	1.2	234	103	AAQ483178	AAQ483178
c 770	c 770	36	1.2	219	70	AW311182	AW311182	sg35e12.y	c 843	36	1.2	235	85	F34689	F34689
c 771	c 771	36	1.2	219	85	F37319	F37319	HSPD35559.H	c 844	36	1.2	236	23	AA239005	AA239005
c 772	c 772	36	1.2	219	86	H84469	H84469	yv86f12.r1	c 845	36	1.2	236	23	AA239005	AA239005
c 773	c 773	36	1.2	220	70	AW312421	AW312421	4218.MARC	c 846	36	1.2	236	40	AI521077	AI521077
c 774	c 774	36	1.2	221	23	AA3270235	AA3270235	va65a07.r	c 847	36	1.2	236	42	AI678494	AI678494
c 775	c 775	36	1.2	221	32	AA918063	AA918063	ol81a12.s	c 848	36	1.2	236	63	AW087987	AW087987
c 776	c 776	36	1.2	221	35	AI144071	AI144071	q163e07.x	c 849	36	1.2	237	29	AA192029	AA192029
c 777	c 777	36	1.2	221	35	AI203594	AI203594	qf57g03.x	c 850	36	1.2	237	23	AA269780	AA269780
c 778	c 778	36	1.2	221	39	AI414351	AI414351	mb64e11.x	c 851	36	1.2	237	42	AI653303	AI653303
c 779	c 779	36	1.2	221	41	AI569953	AI569953	tr90c02.x	c 852	36	1.2	238	21	AA123988	AA123988
c 780	c 780	36	1.2	221	42	AI643858	AI643858	vu10d11.x	c 853	36	1.2	238	39	AI451648	AI451648
c 781	c 781	36	1.2	221	79	AW659288	AW659288	96321.MAR	c 854	36	1.2	238	42	AI641467	AI641467
c 782	c 782	36	1.2	221	86	H83576	H83576	yv81a04.r1	c 855	36	1.2	238	64	AW132015	AW132015
c 783	c 783	36	1.2	222	20	AA017921	AA017921	mh46b08.r	c 856	36	1.2	238	93	AQ078631	AQ078631
c 784	c 784	36	1.2	222	22	AA189894	AA189894	mu55c12.r	c 857	36	1.2	239	37	AI272328	AI272328
c 785	c 785	36	1.2	222	23	AA258131	AA258131	zs35f04.s	c 858	36	1.2	239	39	AI452772	AI452772

859 36 1.2 239 71 AW325910 16971 MAR AW325910 16971 MAR c 932
860 36 1.2 240 20 AA021163 ze6se10.s AA021163 ze6se10.s c 933
861 36 1.2 240 27 AA519067 v126h07.r AA519067 v126h07.r c 934
862 36 1.2 240 47 AL043632 DKFP4341 AL043632 DKFP4341 c 935
863 36 1.2 240 48 AW073934 AW073934 c 936
864 36 1.2 241 22 AA204591 mu59a11.r AA204591 mu59a11.r c 937
865 36 1.2 241 62 AW020414 df09d09.y AW020414 df09d09.y c 938
866 36 1.2 241 63 AW101194 sd75c08.y AW101194 sd75c08.y c 939
867 36 1.2 241 70 AW312799 5109 MARC AW312799 5109 MARC c 940
868 36 1.2 241 120 B51917 CJT-HSP-384 B51917 CJT-HSP-384 c 941
869 36 1.2 242 27 AA472916 vb31c02.r AA472916 vb31c02.r c 942
870 36 1.2 242 28 AA583602 mn61e05.s AA583602 mn61e05.s c 943
871 36 1.2 242 34 AI011695 EST206146 AI011695 EST206146 c 944
872 36 1.2 242 38 AI345463 lb82f01.x AI345463 lb82f01.x c 945
873 36 1.2 242 38 AI382313 ta71c08.x AI382313 ta71c08.x c 946
874 36 1.2 242 63 AW102544 sg60f05.y AW102544 sg60f05.y c 947
875 36 1.2 242 63 AW105459 u53a05.x AW105459 u53a05.x c 948
876 36 1.2 242 79 AW656544 108855 MA AW656544 108855 MA c 949
877 36 1.2 243 26 AA401240 zv63g02.r AA401240 zv63g02.r c 950
878 36 1.2 243 44 AI766785 w188c10.x AI766785 w188c10.x c 951
879 36 1.2 244 31 AA805575 ob44g03.s AA805575 ob44g03.s c 952
880 36 1.2 244 37 AI251769 qu76e02.x AI251769 qu76e02.x c 953
881 36 1.2 244 41 AI545287 f806c03.x AI545287 f806c03.x c 954
882 36 1.2 244 63 AW082599 xc20b09.x AW082599 xc20b09.x c 955
883 36 1.2 244 79 AW620896 sj50a12.y AW620896 sj50a12.y c 956
884 36 1.2 244 80 AW733574 sk75d03.y AW733574 sk75d03.y c 957
885 36 1.2 244 85 H42271 yo63b07.r1 H42271 yo63b07.r1 c 958
886 36 1.2 245 30 AA761917 rz24h03.s AA761917 rz24h03.s c 959
887 36 1.2 245 37 AI289992 qw25e11.x AI289992 qw25e11.x c 960
888 36 1.2 245 70 AW306382 se49h09.y AW306382 se49h09.y c 961
889 36 1.2 245 70 AW311640 s943d06.y AW311640 s943d06.y c 962
890 36 1.2 245 73 AW485189 64429 MAR AW485189 64429 MAR c 963
891 36 1.2 246 23 AA238442 mx92h10.r AA238442 mx92h10.r c 964
892 36 1.2 246 23 AA268648 va43c04.r AA268648 va43c04.r c 965
893 36 1.2 246 30 AA710129 vc45a02.r AA710129 vc45a02.r c 966
894 36 1.2 246 42 AI653918 ty03c05.x AI653918 ty03c05.x c 967
895 36 1.2 246 43 AI696588 tx68b08.x AI696588 tx68b08.x c 968
896 36 1.2 246 47 AW031785 AU031785 AW031785 AU031785 c 969
897 36 1.2 246 70 AW258200 uc31f10.y AW258200 uc31f10.y c 970
898 36 1.2 246 71 AW336324 23232 MAR AW336324 23232 MAR c 971
899 36 1.2 246 71 AW337090 22233 MAR AW337090 22233 MAR c 972
900 36 1.2 247 33 AA936138 om57d02.s AA936138 om57d02.s c 973
901 36 1.2 247 37 AI289937 qv08a03.x AI289937 qv08a03.x c 974
902 36 1.2 247 40 AI522809 fb2f210.x AI522809 fb2f210.x c 975
903 36 1.2 247 43 AI725206 1225 Pt1F AI725206 1225 Pt1F c 976
904 36 1.2 247 48 AU058329 AU058329 AU058329 AU058329 c 977
905 36 1.2 247 63 AW065234 614038B07 AW065234 614038B07 c 978
906 36 1.2 248 30 AA738097 nx11h10.s AA738097 nx11h10.s c 979
907 36 1.2 248 30 AA767127 oa64e09.s AA767127 oa64e09.s c 980
908 36 1.2 248 39 AI445994 tj32c09.x AI445994 tj32c09.x c 981
909 36 1.2 248 47 AU033156 AU033156 AU033156 AU033156 c 982
910 36 1.2 248 72 AW414803 48629 MAR AW414803 48629 MAR c 983
911 36 1.2 248 72 AA430974 71107 MAR AA430974 71107 MAR c 984
912 36 1.2 249 27 AA507581 nh76g03.s AA507581 nh76g03.s c 985
913 36 1.2 249 34 AI068033 E8T209722 AI068033 E8T209722 c 986
914 36 1.2 249 39 AI427562 mm17e06.x AI427562 mm17e06.x c 987
915 36 1.2 249 44 AI783562 tz37c08.x AI783562 tz37c08.x c 988
916 36 1.2 249 70 AW311129 9933c11.y AW311129 9933c11.y c 989
917 36 1.2 249 71 AW359405 45979 MAR AW359405 45979 MAR c 990
918 36 1.2 249 72 AA415599 49989 MAR AA415599 49989 MAR c 991
919 36 1.2 249 72 AA427866 64424 MAR AA427866 64424 MAR c 992
920 36 1.2 250 22 AA147342 zc39b03.s AA147342 zc39b03.s c 993
921 36 1.2 250 32 AA866865 u06c05.r AA866865 u06c05.r c 994
922 36 1.2 251 29 AA687139 nv63e05.s AA687139 nv63e05.s c 995
923 36 1.2 251 30 AA734985 v517g07.r AA734985 v517g07.r c 996
924 36 1.2 251 31 AA805063 ob86a12.s AA805063 ob86a12.s c 997
925 36 1.2 251 31 AA811218 of77h07.s AA811218 of77h07.s c 998
926 36 1.2 251 32 AA877935 o577h07.s AA877935 o577h07.s c 999
927 36 1.2 251 38 AI344947 tb01c03.x AI344947 tb01c03.x c 1000
928 36 1.2 251 41 AI553392 vw51f11.x AI553392 vw51f11.x c 997
929 36 1.2 251 41 AI597465 v164g11.x AI597465 v164g11.x c 998
930 36 1.2 251 63 AW084330 xc55c06.x AW084330 xc55c06.x c 999
931 36 1.2 252 23 AA230854 mw14b02.r AA230854 mw14b02.r c 1000

ALIGNMENTS


```

QY 818 ggaagactggaagatcggttccaactgtgcaaccagacaacaaatcagactg 869
|||||
Db 257 GGAAGGACTGGAAGATCGGCTTCCAACTGTGCAACCAAGACAAATCAGACTG 308
|||||

RESULT 5
AA980401/c
LOCUS
DEFINITION
ua52g11.r1 Stratagene mouse lung 937302 Mus musculus cDNA clone
IMAGE:1350404 5' similar to SW:SCAA_RAT P37089 AMILORIDE-SENSITIVE
SODIUM CHANNEL ALPHA-SUBUNIT ; mRNA sequence.
ACCESSION
AA980401
VERSION
AA980401.1 GI:3158937
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 442)
REFERENCE
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
On Apr 7, 1998 this sequence version replaced gi:3035617.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:699196
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 439.
Location/Qualifiers
1. .442
/organism="Mus musculus"
/strain="C57BL/6 x CBA"
/db_xref="taxon:10090"
/clone_lib="Stratagene mouse lung 937302"
/sex="female"
/tissue_type="lung"
/dev_stage="6-8 month old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. 6-8 month old female lung and 1.5 year old male
lung were source of mRNA. Average insert size: 1.5 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTGTGTGTGTGTGTGT 3'
1 others
BASE COUNT 112 a 114 c 150 g 65 t
ORIGIN
Query Match 1.4%; Score 45; DB 33; Length 442;
Best Local Similarity 100.0%; Pred. No. 8.8e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2061 gacccctccctgagccctgagccctccacctgctatgctac 2105
|||||
Db 213 GACCCCTCCCTGGCCCTGACAGCCCTCCACCTGCCTATGCTAC 169
|||||

RESULT 6
AL040100/c
LOCUS
DEFINITION
AL040100.1
ACCESSION
AL040100
VERSION
AL040100.1 GI:5409067
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157)
REFERENCE
Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Koehler, et al.)
Unpublished (1999)
On May 18, 1998 this sequence version replaced gi:3138571.
Contact: Koehler K
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
r1 sequence also available.
This clone (DKFZp434B0913) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434B0913"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
1 others
BASE COUNT 38 a 14 c 25 g 79 t
ORIGIN
Query Match 1.3%; Score 41; DB 47; Length 157;
Best Local Similarity 100.0%; Pred. No. 7.8e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 54 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 14
|||||

RESULT 7
AW278153
LOCUS
DEFINITION
sf40408.y1 Gm-cl009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl009-2440 5', mRNA sequence.
ACCESSION
AW278153
VERSION
AW278153.1 GI:6666694
KEYWORDS
EST.
SOURCE
soybean.
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales;
Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 157)
REFERENCE
Shoenaker, R., Keim, P., Vodkin, L., Erpelding, J., Corvelli, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

```

COMMENT

On Oct 8, 1998 this sequence version replaced gi:3727461.
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
 info@genomesystems.com web site: www.genomesystems.com
 Seq primer: -40R from Gibco
 High quality sequence stop: 132.

FEATURES

source
 1. .157
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl009-2440"
 /clone_lib="Gm-cl009"
 /lab_host="XL10-Gold"
 /note="vector: pBluescript II XR; Site_1: EcoRI; Site_2:
 XhoI; The mRNA was isolated from entire roots of
 2-month-old 'Williams' plants that were greenhouse grown
 in 5-gallon pots. To suppress nodulation, Black Gold
 All-Purpose potting soil was supplemented with: 0.36g/L
 available phosphoric acid (P205), 20mg/L urea N, 0.16g/L
 S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn,
 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following
 nutrients in a slow-release form (Osmocote): 0.165g/L
 ammoniac N, 0.185g/L nitrate N, 0.35g/L available
 phosphoric acid, and 0.35g/L soluble potash. No nodules
 were visible on the roots at harvest. Stratagene's cDNA
 Synthesis Kit (catalog #200401) was used to synthesize the
 cDNA. First-strand synthesis was performed with 5-methyl
 dCTP, hence the ligated cDNA is hemimethylated.
 Stratagene's first-strand synthesis primer was used
 [GACAGAGACAGACAGAGAACTACTGCGAG(T)-18]. After
 second-strand synthesis is, the cDNA ends were 'polished'
 with clone Pfu DNA polymerase, ligated to EcoRI adapters,
 and phosphorylated. The XhoI site within the first-strand
 synthesis primer was restricted by digestion with XhoI;
 all XhoI sites in the cDNA would be protected by their
 hemimethylated status. The cDNA constructs were
 size-fractionated with a 400bp cutoff, using a SizeSep 400
 spin column from Pharmacia. The column eluent was then
 ligated into Stratagene's pBluescript II XR Predigested
 vector (pBluescript II SK(+)) that had been digested with
 EcoRI and XhoI, and phosphorylated). Both the white and
 blue colonies appear to contain recombinant plasmids with
 cDNA inserts. This library was constructed by Dr. Paul
 Keim and Dr. Virginia Coryell."

BASE COUNT

99 a 21 c 12 g 25 t

Query Match 1.3%; Score 41; DB 70; Length 157;
 Best Local Similarity 100.0%; Pred. No. 7.8e-09;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||||||||||||||||||||||||||||||||||||
 Db 69 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 109

RESULT 8

AU073712 240 bp mRNA EST 24-JUN-1999
 LOCUS AU073712 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
 DEFINITION AU073712 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
 accession AU073712
 VERSION AU073712.1 GI:5180133

KEYWORDS

SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum
 REFERENCE Eukaryote; Dictyostellida; Dictyostelium.
 AUTHORS Urushihara, H.
 TITLE Developmental cDNA in Dictyostelium discoideum (1999)
 JOURNAL Unpublished (1999)
 COMMENT On Jul 8, 1999 this sequence version replaced gi:5422501.
 Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
 Email: d402hu@sakura.cc.tsukuba.ac.jp
 PROJECT = Dictyostelium discoideum cDNA project in Japan.

FEATURES

source
 1. .240
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="SSI133"
 /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
 /dev_stage="slug" 30 g 85 t
 BASE COUNT 108 a 17 c 30 g 85 t
 ORIGIN

Query Match 1.3%; Score 41; DB 48; Length 240;
 Best Local Similarity 100.0%; Pred. No. 7.7e-09;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||||||||||||||||||||||||||||||||||||
 Db 47 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 87

RESULT 9

AU065863 262 bp mRNA EST 30-NOV-1999
 LOCUS AU065863 Rice green shoot Oryza sativa cDNA clone SI4374_1A, mRNA
 DEFINITION AU065863 sequence.
 ACCESSION AU065863
 VERSION AU065863.2 GI:6481240
 KEYWORDS EST.
 SOURCE Oryza sativa.

ORGANISM

Oryza sativa
 Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

REFERENCE

1 (bases 1 to 262)
 Yamamoto, K. and Sasaki, T.
 Rice cDNA from green shoot
 Unpublished (1996)

JOURNAL

On Jun 2, 1999 this sequence version replaced gi:4968959.
 Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program
 2-1-2 Kannondai, Tsukuba
 Ibaraki,
 Japan 305
 Tel: 0298-38-7441
 Fax: 0298-38-7468
 Email: tsasaki@abr.affrc.go.jp
 PROJECT = 'RGP'

COMMENT

Sequence updated (24-Nov-1999).
 Location/Qualifiers
 1. .262
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone="SI4374_1A"
 /note="Green shoot (8 days old)"

FEATURES

source
 1. .262
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone="SI4374_1A"
 /note="Green shoot (8 days old)"
 BASE COUNT 94 a 39 c 52 g 77 t

ORIGIN

Query Match 1.3%; Score 41; DB 48; Length 262;
Best Local Similarity 100.0%; Pred. No. 7.7e-09; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 222 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 262

RESULT 10

AF034791 1367 bp mRNA EST 22-APR-1998

LOCUS AF034791 Felis catus partial mRNA Felis catus cDNA similar to
DEFINITION glyceraldehyde 3-phosphate dehydrogenase, mRNA sequence.

ACCESSION AF034791

VERSION AF034791.1 GI:2654165

KEYWORDS EST.

SOURCE cat.

ORGANISM Felis catus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

REFERENCE 1 (bases 1 to 1367)

AUTHORS Helps,C.R. and Harbour,D.A.

TITLE Felis catus glyceraldehyde 3-phosphate dehydrogenase mRNA

JOURNAL Unpublished (1997)

COMMENT On Apr 7, 1998 this sequence version replaced gi:3036550.

Contact: Helps, Chris R.

Clinical and Veterinary Science

University of Bristol

Bristol BS18 7DU, UK

Email: c.r.helps@bris.ac.uk.

Location/Qualifiers

1. .1367

/organism="Felis catus"

/db_xref="taxon:9685"

/clone_lib="Felis catus partial mRNA"

/cell_line="FL4"

BASE COUNT 325 a 317 c 376 g 276 t 73 others

ORIGIN

Query Match 1.3%; Score 41; DB 33; Length 1367;
Best Local Similarity 100.0%; Pred. No. 7.4e-09; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 1302 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1342

RESULT 11

AW313563 210 bp mRNA EST 24-JAN-2000

LOCUS AW313563 8840 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION AW313563

ACCESSION AW313563

VERSION AW313563.1 GI:6742783

KEYWORDS EST.

SOURCE Bos taurus.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

REFERENCE 1 (bases 1 to 210)

AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,

Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and

Keele,J.W.

TITLE Design and use of four pooled tissue normalized cDNA libraries for

EST discovery in cattle

Unpublished (2000)

COMMENT On Jul 7, 1999 this sequence version replaced gi:5410190.

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 20

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 5 row: E column: 22

Seq primer: ATTTAGTGACACTATAG.

FEATURES

source

1. .210

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 1BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: PCMV SPORT6; Site.1: XbaI; Site.2: XhoI;

Library made from pooled tissue from lymph node, ovary,

fat, hypothalamus, and pituitary."

BASE COUNT 89 a 42 c 35 g 44 t

ORIGIN

Query Match 1.3%; Score 40; DB 70; Length 210;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tcctcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 157 TCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 196

RESULT 12

AW657118 224 bp mRNA EST 05-APR-2000

LOCUS AW657118 109764 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION AW657118

ACCESSION AW657118

VERSION AW657118.1 GI:7422944

KEYWORDS EST.

SOURCE Bos taurus.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

REFERENCE 1 (bases 1 to 224)

AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,

Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and

Keele,J.W.

TITLE Design and use of four pooled tissue normalized cDNA libraries for

EST discovery in cattle

Unpublished (2000)

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 99 row: H column: 21

Seq primer: ATTTAGTGACACTATAG.

Location/Qualifiers

1. .224

/organism="Bos taurus"

```

/db_xref="taxon:9913"
/clone_lib="MARC IBOV"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
BASE COUNT      84 a  49 c  51 g  40 t
ORIGIN

```

```

Query Match      1.3%; Score 40; DB 79; Length 224;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 97  TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 136

```

```

RESULT 13
LOCUS      C98601      275 bp      mRNA      EST      19-OCT-1998
DEFINITION C98601 Rice panicle at flowering stage Oryza sativa cDNA clone
E0453_62, mRNA sequence.
ACCESSION C98601.1 GI:3761353
KEYWORDS   EST.
SOURCE     Oryza sativa.
ORGANISM   Oryza sativa
REFERENCE  1 (bases 1 to 275)
AUTHORS    Sasaki,T. and Yamamoto,K.
TITLE      Rice cDNA from panicle at flowering stage
JOURNAL    Unpublished (1996)
COMMENT    On Jun 22, 1998 this sequence version replaced gi:3247562.
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program
2-1-2 Kannondai, Tsukuba
Ibaraki,
Japan 305
Tel: 0298-38-7441
Fax: 0298-38-7468
Email: tsasaki@abr.affrc.go.jp
PROJECT - "RGP".

```

```

FEATURES
source
Location/Qualifiers
1..275
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone_lib="E0453_62"
/dev_stage="flowering stage"
/notes="Organ: panicle; Rice cDNA from panicle at flowering
stage"
BASE COUNT      113 a  44 c  42 g  70 t  6 others
ORIGIN

```

```

Query Match      1.3%; Score 40; DB 81; Length 275;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 225  TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 264

```

```

RESULT 14
LOCUS      AI253246/c      449 bp      mRNA      EST      05-NOV-1998

```

```

DEFINITION
ACCESSION      AI253246
VERSION        AI253246.1
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens

```

```

REFERENCE
AUTHORS        Robert Strausberg, Ph.D.
TITLE          Contact: Robert Strausberg@nih.gov
JOURNAL        Unpublished (1997)
COMMENT        Tel: (301) 496-1550
                Email: Robert.Strausberg@nih.gov
                unknown library type
                Seq primer: -40UP from Gibco
                High quality sequence stop: 436.
                Location/Qualifiers
                1..449

```

```

FEATURES
source
Location/Qualifiers
1..449
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP_Kid1"
/lab_host="DH10B"
/notes="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaudo."
BASE COUNT      145 a  57 c  63 g  184 t
ORIGIN

```

```

Query Match      1.3%; Score 40; DB 37; Length 449;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 41  TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

```

```

RESULT 15
LOCUS      AI320733/c      79 bp      mRNA      EST      18-DEC-1998
DEFINITION AI320733 Neurospora crassa morning cDNA library Neurospora crassa
cDNA clone c9f03nm 5', mRNA sequence.
ACCESSION      AI320733
VERSION        AI320733.1
KEYWORDS       EST.
SOURCE         Neurospora crassa.
ORGANISM       Neurospora crassa

```

```

REFERENCE
AUTHORS        Zhu,H., Lai,H., Kupfer,D., Dunlap,J.C. and Roe,B.A.
TITLE          Two Neurospora crassa EST Databases
JOURNAL        Unpublished (1998)
COMMENT        On May 18, 1998 this sequence version replaced gi:3138462.
                Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
                Department of Chemistry and Biochemistry
                Advanced Center for Genome Technology, University of Oklahoma
                620 Parrington Oval, Norman, OK 73019, USA
                Tel: 405 325 4912
                Fax: 405 325 7762

```

Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
 Seq primer: Universal Forward primer
 High quality sequence stop: 56.
 Location/Qualifiers

FEATURES

source

```
1..79
/organism="Neurospora crassa"
/strain="bd, frq7 A"
/db_xref="taxon:5141"
/clone="c9f03nm"
/clone_lib="Neurospora crassa morning cDNA library"
/tissue_type="tissue harvested following 22hr growth in dark"
/note="Vector: pBluescript SK-; Site1: XbaI; Site2: EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096,1996. 5' end of cDNA cloned into XbaI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"
```

BASE COUNT

ORIGIN

Query Match 1.3%; Score 39; DB 38; Length 79;
 Best Local Similarity 100.0%; Pred. No. 7.4e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 40 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

RESULT 16

LOCUS

AI946912 139 bp mRNA EST 17-AUG-1999
 bs32h08.y1 Drosophila melanogaster adult testis library Drosophila melanogaster cDNA clone bs32h08 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 139)

Andrews, J., Bouffard, G. and Oliver, B.

Drosophila melanogaster testis expressed sequence tags

Unpublished (1999)

Contact: Brian Oliver

Laboratory of Cellular and Developmental Biology

NIDDK, National Institutes of Health

6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA

Fax: (301) 496 5239

Email: oliver@helix.nih.gov,

http://www.niddk.nih.gov/intram/people/boliver.htm

Tissue isolation and library construction performed at the National Institute of Diabetes and Digestive and Kidney Diseases, NIH (see http://www.niddk.nih.gov/intram/people/boliver.htm). DNA sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC; see http://www.nisc.nih.gov). Chromatogram data were analyzed and evaluated for high quality using the ted program (Gleeson T and Hillier L, 1991).

Plate: 32 row: h column: 08

Seq primer: M13RP1 reverse primer (ABI).

Location/Qualifiers

1..139

/organism="Drosophila melanogaster"

/strain="v[*] w[67cl]/Y"

/db_xref="taxon:7227"

/clone="bs32h08"

/clone_lib="Drosophila melanogaster adult testis library"

/sex="male"

/dev_stage="1-5 day adult"

FEATURES

source

FEATURES

source

1..161

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2176648"

/clone_lib="NCI_CGAP_Ut2"

/tissue_type="moderately-differentiated endometrial

adenocarcinoma, 3 pooled tumors"

/lab_host="DH10B"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site1: SalI;

Site2: NotI; Cloned unidirectionally. Primer: Oligo df.

Average insert size 1.85 kb. Life Technologies catalog #:

11539-012"

34 a 10 c 35 g 82 t

BASE COUNT

ORIGIN

/lab_host="SOLR (Stratagene)"
 /note="Organ: testis; Vector: pBluescript SK (Stratagene); Site1: EcoR I; Site2: Xho I; Testes dissected from 1-5 day adult y[*] w[67cl]/Y males raised at 25OC. RNA isolated using Trizol (Life Technologies) and a single round of Poly(A)+ selection using Oligotex (Qiagen). cDNA library constructed using Stratagene ZAP-cDNA syntesis kit. Oligo dt-primed, size fractionated -1-6 kb, and directionally cloned at EcoRI and XhoI in Uni-ZAP XR. Following a single round of amplification pBluescript SK phagemids were mass excised. A distribution channel for clones is being sought, but not currently available. Requests for clones cannot be honored."

BASE COUNT 74 a 30 c 26 g 7 t 2 others

ORIGIN

Query Match 1.3%; Score 39; DB 46; Length 139;

Best Local Similarity 100.0%; Pred. No. 7.3e-08;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 69 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 107

RESULT 17

LOCUS

AI540624 161 bp mRNA EST 14-APR-1999
 tn88e09.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2176648 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 161)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On May 18, 1998 this sequence version replaced gi:3136848.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/dbp/image/image.html

Insert Length: 1843 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 116.

Location/Qualifiers

1..161

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2176648"

/clone_lib="NCI_CGAP_Ut2"

/tissue_type="moderately-differentiated endometrial

adenocarcinoma, 3 pooled tumors"

/lab_host="DH10B"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site1: SalI;

Site2: NotI; Cloned unidirectionally. Primer: Oligo df.

Average insert size 1.85 kb. Life Technologies catalog #:

11539-012"

34 a 10 c 35 g 82 t

BASE COUNT

ORIGIN

Query Match 1.3%; Score 39; DB 41; Length 161;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||
 Db 39 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 18
 AA231085 167 bp mRNA EST 26-FEB-1997
 mw11e1.rl Soares mouse 3NME12 5 Mus musculus cDNA clone
 IMAGE:670412 5', mRNA sequence.

ACCESSION AA231085
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE 1 (bases 1 to 167)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:41016

Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 155.

FEATURES
 Source
 1. .167
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:670412"
 /clone_lib="Soares mouse 3NME12 5"
 /sex="unknown"
 /tissue_type="fetus"
 /dev_stage="12.5dpc total fetus"
 /lab_host="DH10B"
 /note="Organ: whole fetus; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' GTTTCACCATCTCAAGTCGGAGCGCGCTATTATTTTATTTTATTTT 3'], on total mouse RNA [provided by Minoru Ko, Wayne State Univ.]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 97 a 20 c 18 g 32 t

ORIGIN
 Query Match 1.3%; Score 39; DB 23; Length 167;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

RESULT 20
AW477617 190 bp mRNA EST 24-FEB-2000
LOCUS
DEFINITION 15122 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AW477617
VERSION AW477617.1 GI:7047723
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 190)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT On Jan 6, 2000 this sequence version replaced gi:6674418.
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCAGCAGC
Plate: 7 row: N column: 9
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .190
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 79 a 53 c 20 g 38 t
ORIGIN
Query Match 1.3%; Score 39; DB 73; Length 190;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 140 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 178
|||||
RESULT 21
AW172002 207 bp mRNA EST 15-NOV-1999
LOCUS 618047c10.y1 618 - Inbred Tassel cDNA Library Zea mays cDNA, mRNA
DEFINITION sequence.
ACCESSION AW172002
VERSION AW172002.1 GI:6431798
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
REFERENCE 1 (bases 1 to 207)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University

JOURNAL
COMMENT

Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 618047 row: C column: 10.
Location/Qualifiers
1. .207
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="618 - Inbred Tassel cDNA Library"
/tissue_type="tassel"
/dev_stage="tassel length from 0.1 to 2.5 cm"
/lab_host="XLOU"
/note="Organ: tassel; Vector: pAD-GAL4-2.1 (Hybrizap);
Inbred tassel library from Schmidt lab"

FEATURES
source

BASE COUNT 83 a 29 c 35 g 60 t
ORIGIN

Query Match 1.3%; Score 39; DB 64; Length 207;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 151 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 189
|||||

RESULT 22

AW357288 229 bp mRNA EST 03-FEB-2000
LOCUS 40265 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION AW357288
ACCESSION AW357288.1 GI:6861294
VERSION
KEYWORDS EST.
SOURCE Bos taurus.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 229)
AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and
Keele,J.W.

REFERENCE
AUTHORS

Design and use of four pooled tissue normalized cDNA libraries for
EST discovery in cattle
Unpublished (2000)

TITLE

On Apr 7, 1998 this sequence version replaced gi:3034614.

JOURNAL

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCAGCAGC
Plate: 18 row: G column: 6
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .229
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"

FEATURES
source

/lab_host="DH108"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."
 96 a 51 c 49 g 33 t

BASE COUNT
 ORIGIN

Query Match 1.3%; Score 39; DB 71; Length 229;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 157 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 195

RESULT 23
 AA832425/c
 LOCUS
 DEFINITION
 mRNA sequence.
 AA832425
 VERSION
 AA832425.1 GI:2905524
 EST
 SOURCE
 HOMO
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 245)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Jan 19, 1998 this sequence version replaced gi:2286298.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550

AA832425 245 bp mRNA EST 25-MAR-1998
 OC99d12.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1357847 3',
 mRNA sequence.
 AA832425
 VERSION
 AA832425.1 GI:2905524
 EST
 SOURCE
 HOMO
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 245)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Jan 19, 1998 this sequence version replaced gi:2286298.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/brp/image/image.html
 Insert Length: 992 Std Error: 0.00
 Seq primer: -40m13 fwd. Et from Amersham
 High quality sequence stop: 227.
 Location/Qualifiers
 1. .245
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1357847"
 /clone_lib="NCI-CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH108"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 42 a 58 c 64 g 81 t

BASE COUNT
 source

ORIGIN

Query Match 1.3%; Score 39; DB 31; Length 245;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 40 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

RESULT 24
 AW056177
 LOCUS
 DEFINITION
 mRNA sequence.
 AW056177
 VERSION
 AW056177.1 GI:5928885
 EST
 SOURCE
 Zea mays.
 Zea mays.
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
 1 (bases 1 to 281)
 Walbot.V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 On Jun 15, 1998 this sequence version replaced gi:3222524.
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660005 row: C column: 01.

AW056177 281 bp mRNA EST 27-SEP-1999
 660005C01.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
 mRNA sequence.
 AW056177
 VERSION
 AW056177.1 GI:5928885
 EST
 SOURCE
 Zea mays.
 Zea mays.
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
 1 (bases 1 to 281)
 Walbot.V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 On Jun 15, 1998 this sequence version replaced gi:3222524.
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660005 row: C column: 01.

Location/Qualifiers
 1. .281
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="XLOUR"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."
 BASE COUNT 133 a 37 c 33 g 78 t

BASE COUNT
 ORIGIN

Query Match 1.3%; Score 39; DB 63; Length 281;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 236 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274

RESULT 25
 AW479416
 LOCUS
 DEFINITION
 mRNA
 25195 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
 AW479416
 ACCESSION
 AW479416.1 GI:7049522
 VERSION
 EST.
 SOURCE
 Bos taurus.
 Bos taurus.
 ORGANISM

BASE COUNT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 307)
Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and Keele,J.W.

TITLE
Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle

JOURNAL
Unpublished (2000)

COMMENT
On Apr 7, 1998 this sequence version replaced gi:3034581.

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 17 row: P column: 12

Seq primer: ATTAGGTGACACTATAG.

FEATURES
source

1..307
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."

BASE COUNT 118 a 65 c 39 g 85 t
ORIGIN

Query Match 1.3%; Score 39; DB 73; Length 307;

Best Local Similarity 100.0%; Pred. No. 7.1e-08;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 258 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 296

RESULT 26

AA237390

LOCUS

DEFINITION AA237390.1 318 bp mRNA EST 03-MAR-1997
mx18b02.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:680523 5' similar to gb:M14058 COMPLEMENT CIR COMPONENT PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 318)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,J.S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheilenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

TITLE On Oct 30, 1996 this sequence version replaced gi:1656971.

JOURNAL Contact: Marra M/Mouse EST Project

COMMENT WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:420227

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 267.

FEATURES
source

Location/Qualifiers

1..318

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_lib="IMAGE:680523"

/clone_lib="Soares mouse NML"

/tissue_type="Liver"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGGGAATCTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 128 a 45 c 68 g 77 t
ORIGIN

Query Match 1.3%; Score 39; DB 23; Length 318;

Best Local Similarity 100.0%; Pred. No. 7.1e-08;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 275 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 313

RESULT 27

AU082521

LOCUS

DEFINITION AU082521 327 bp mRNA EST 17-FEB-2000
E30614, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

REFERENCE 1 (bases 1 to 327)

AUTHORS Sasaki,T. and Yamamoto,K.

TITLE Rice cDNA from panicle (2000)

JOURNAL Unpublished (2000)

COMMENT Contact: Takuii Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program

2-1-2 Kannondai,Tsukuba

Ibaraki,

Japan 305

Tel: 0298-38-7441

Fax: 0298-38-7468

Email: tsasaki@abr.affrc.go.jp

PROJECT = 'RGP'.

E30614_6Z.

FEATURES

source

Location/Qualifiers

1..327

/organism="Oryza sativa"

/strain="Nipponbare"

/db_xref="taxon:4530"

/clone_lib="E30614"

/clone_lib="Rice panicle shorter than 3cm"

LOCUS	AW394641	341 bp	mRNA	EST	07-FEB-2000
DEFINITION	Sh05a07.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1016-3949.5, similar to TR:O65075 O65075 HYPOTHETICAL 30.9 KD PROTEIN ; mRNA sequence.				
ACCESSION	AW394641				

VERSION
KEYWORDS
SOURCE
ORGANISM

AW394641.1 GI:6913111
EST.
soybean.
Glycine max

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales;
Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 341)
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project

TITLE
JOURNAL
COMMENT

Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com

FEATURES
source

Seq primer: -40RP from Gibco.
Location/Qualifiers
1..341
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-3949"
/clone_lib="Gm-cl016"
/tissue_type="immature flowers of field grown plants"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site.1: EcoRI; Site.2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into XL10-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpelding."

BASE COUNT 122 a 61 c 60 g 98 t
ORIGIN

Query Match 1.3%; Score 39; DB 71; Length 341;
Best Local Similarity 100.0%; Pred. No. 7.le-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3076 gtccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3114
|||||
Db 303 GTTCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341

RESULT 31
AI529904

LOCUS AI529904 376 bp mRNA EST 18-MAR-1999
DEFINITION u183h01.v1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1889041 5' similar to SW:OATP_RAT P46720 SODIUM-INDEPENDENT
ORGANIC ANION TRANSPORTER ;, mRNA sequence.
ACCESSION AI529904
VERSION AI529904.1 GI:4444039
KEYWORDS EST.

SOURCE
ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 376)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999

TITLE
JOURNAL
COMMENT

Unpublished (1999)
On Jan 19, 1998 this sequence version replaced gi:2286451.
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:973365

Seq primer: custom primer used
High quality sequence stop: 361.

FEATURES
source

Location/Qualifiers
1..376
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889041"
/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site.1: DraIII
(CACGTGTG); Site.2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTGCG], digested
and cloned into distinct draIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTCTCTAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGACACA."

BASE COUNT 147 a 61 c 70 g 98 t
ORIGIN

Query Match 1.3%; Score 39; DB 40; Length 376;
Best Local Similarity 100.0%; Pred. No. 7.le-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 309 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 347

RESULT 32
D22652

LOCUS D22652 380 bp mRNA EST 08-JUL-1999
DEFINITION RICC0739B Rice callus Oryza sativa cDNA clone C0739_52, mRNA
sequence.
ACCESSION D22652
VERSION D22652.1 GI:425969
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
REFERENCE 1 (bases 1 to 380)


```

Query Match      1.3%; Score 39; DB 34; Length 427;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 43 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 35
AA610049/c
LOCUS      444 bp mRNA EST
DEFINITION af18h01.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1032049
3', mRNA sequence.
ACCESSION AA610049
VERSION AA610049.1 GI:2458477
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2284987.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 956 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 262.
FEATURES
Location/Qualifiers
source
1..444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1032049"
/tissue="male"
/lab_host="Soares_testis_NHT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5',
TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 89 a 103 c 101 g 151 t
ORIGIN
Query Match      1.3%; Score 39; DB 28; Length 444;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 50 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 36
AA731682/c
LOCUS      450 bp mRNA EST
DEFINITION nw58e10.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1250826 3'
similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA731682
VERSION AA731682
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 542 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 363.
FEATURES
Location/Qualifiers
source
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1250826"
/tissue="male"
/lab_host="NCI_CGAP_GCB1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI). Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT-3',
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 95 a 85 c 84 g 184 t 2 others
ORIGIN
Query Match      1.3%; Score 39; DB 30; Length 450;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 44 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6

RESULT 37
AW249071
LOCUS      475 bp mRNA EST
DEFINITION AW249071.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820891 5',
mRNA sequence.
ACCESSION AW249071
VERSION AW249071.1 GI:6592064

```

KEYWORDS SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 475)
 AUTHORS NIH-MGC <http://www.ncbi.nlm.nih.gov/MGC/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT On Jul 7, 1999 this sequence version replaced gi:5866355.
 Other_ESTs: 2820891.3prime
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: DCTD/DRP cDNA Library Preparation: Ling
 Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
 Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
 project Clone Distribution: MGC clone distribution information can
 be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
 Scores: PHRED from University of Washington Genome Center
 Trimming: cross_match from University of Washington Genome Center
 PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
 Drosophila Genome Project. University of Washington Genome Center:
<http://www.genome.washington.edu>
 Plate: LLCW5 row: G column: 4
 High quality sequence stop: 397.
 Location/Qualifiers
 1..475
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2820891"
 /clone_lib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G).. Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 101 a 135 c 127 g 112 t
 BASE COUNT
 ORIGIN
 Query Match 1.3%; Score 39; DB 70; Length 475;
 Best Local Similarity 100.0%; Pred. No. 7.1e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1800 cctctgtccaacctgggcagcagtgagcctgtggtt 1838
 ||||||||||||||||||||||||||||||||||||
 Db 306 CCTCTGCTCCAACTGGCAGGCGAGTGGAGCTGTGTT 344
 RESULT 38
 A1018664/c 482 bp mRNA EST 27-AUG-1998
 LOCUS ov65e05.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1642208
 DEFINITION 3' similar to TR:Q63797 Q63797 PROTEASOME ACTIVATOR RPA28 SUBUNIT
 ALPHA.; mRNA sequence.
 ACCESSION A1018664
 VERSION A1018664.1 GI:3232462
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 482)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index
 Unpublished (1997)
 On May 18, 1998 this sequence version replaced gi:3137113.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 871 Std Error: 0.00
 Seg primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 314.
 Location/Qualifiers
 1..482
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1642208"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer {5'
 TGTTCACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTT 3'}.
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 111 a 106 c 111 g 154 t
 BASE COUNT
 ORIGIN
 Query Match 1.3%; Score 39; DB 34; Length 482;
 Best Local Similarity 100.0%; Pred. No. 7.1e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 ||||||||||||||||||||||||||||||||||||
 Db 44 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6
 RESULT 39
 AA393950 499 bp mRNA EST 12-AUG-1997
 LOCUS zt78a10.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:728442
 DEFINITION 5' similar to gb:U29007_cds1 AMILORIDE-SENSITIVE SODIUM CHANNEL
 ALPHA-SUBUNIT (HUMAN);, mRNA sequence.
 ACCESSION AA393950
 VERSION AA393950
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 499)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Scheilenberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 TITLE Unpublished (1997)
 JOURNAL Contact: Wilson RK
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1637 Std Error: 0.00
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 459.

FEATURES

source

```

Location/Qualifiers
1..499
/organism="Homo sapiens"
/db_xref="GDB:5925355"
/db_xref="taxon:9606"
/clone_image="728442"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      111 a 138 c 127 g 121 t
ORIGIN
Query Match      1.3%; Score 39; DB 25; Length 499;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccacacctgggcagccagtcggcgctgtggtt 1838
|||||
DB 372 CCTCCTGTCTCAACCTGGCGCAGTCAGTGGAGCCTGTGGTT 410
|||||

RESULT 40
AA808304/c
LOCUS
DEFINITION
oc4if12.sl NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1352303 3'
similar to TR:Q14012 Q14012 CAM KINASE I.; mRNA sequence.
AA808304
AA808304.1 GI:2877710
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 506)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1141 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 199.
Location/Qualifiers
1..506
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1367973"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1352303"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      113 a 113 c 142 g 138 t
ORIGIN
Query Match      1.3%; Score 39; DB 31; Length 506;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 49 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11
|||||

RESULT 41
AA810286/c
LOCUS
DEFINITION
od14h11.sl NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1367973 3'
similar to SW:IGUP_HUMAN Q06323 INTERPERON GAMMA UP-REGULATED
I-5111 PROTEIN PRECURSOR.; mRNA sequence.
AA810286
AA810286.1 GI:2879645
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 510)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Dec 18, 1997 this sequence version replaced gi:2340745.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 934 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 227.
Location/Qualifiers
1..510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1367973"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"

```

FEATURES

source

/lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 119 a 115 c 114 g 162 t
 ORIGIN
 Query Match 1.3%; Score 39; DB 31; Length 510;
 Best Local Similarity 100.0%; Pred. No. 7.1e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3079 ccatacaaaaaa 519 bp mRNA EST 04-NOV-1998
 ||||||| 11 NCI-CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1867362 3'
 Db 47 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9
 GENE, COMPLETE CDS, AND OS9 ;, mRNA sequence.

RESULT 42
 AI241203/C
 LOCUS
 DEFINITION
 qj96g10.x1 NCI-CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1867362 3'
 similar to TR:000579 000579 COSMID 6E5 CDK4, SAS and KIAA0167
 GENES, COMPLETE CDS, AND OS9 ;, mRNA sequence.
 ACCESSION AI241203
 VERSION AI241203.1 GI:3836600
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 519)
 REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On May 7, 1998 this sequence version replaced gi:3119061.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbrrp/image/image.html
 Seq primer: -400P from Gibco
 High quality sequence stop: 235.
 Location/Qualifiers
 1. .519
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1867362"
 /clone_lib="NCI-CGAP_Kid3"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer,
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pT7T3 vector. mRNA
 source: 2 pooled kidneys. Library went through one round

of normalization. Library constructed by Bento Soares and
 M. Fatima Bonaldo."
 BASE COUNT 116 a 123 c 117 g 163 t
 ORIGIN
 Query Match 1.3%; Score 39; DB 37; Length 519;
 Best Local Similarity 100.0%; Pred. No. 7.1e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3079 ccatacaaaaaa 520 bp mRNA EST 01-MAY-1998
 ||||||| 11 Apple young fruit cDNA library Malus x domestica cDNA
 Db 42 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4
 clone af148, mRNA sequence.

ACCESSION AT000142
 VERSION AT000142.1 GI:3088048
 KEYWORDS
 SOURCE apple tree.
 ORGANISM Malus x domestica
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; Rosidae; eurossids I; Rosales;
 Rosaceae; Malus.
 1 (bases 1 to 520)
 REFERENCE
 AUTHORS Sung,S.-K., Jeong,D.-H., Nam,J., Kim,S.-H., Kim,S.-R. and An,G.
 TITLE Expressed Sequence Tags of Fruits, Peels, and Carpels and Analysis
 of mRNAs Expression Levels of the Tagged cDNAs of Fruits from the
 Fuji Apple
 JOURNAL Unpublished (1998)
 COMMENT On Apr 7, 1998 this sequence version replaced gi:3036618.
 Contact: Gynheung An
 Department of Life Science
 Pohang University of Science and Technology
 San 31 Hyojadong, Pohang Kyungbuk 790-784, Republic of Korea
 Tel: 82-562-279-2176
 Fax: 82-562-279-2199
 Email: genean@postech.ac.kr
 Submitted through BRIC (Biological Research Information Center) of
 Korea URL: http://bric.postech.ac.kr/.
 Location/Qualifiers
 1. .520
 /organism="Malus x domestica"
 /cultivar="Fuji"
 /db_xref="taxon:3750"
 /clone="af148"
 /clone_lib="Apple young fruit cDNA library"
 /dev_stage="7-10 days after anthesis"
 /note="Vector: Lambda vector Unizap II; initial
 pfu:7x10⁶; average insert size: 0.85"
 BASE COUNT 168 a 61 c 160 g 129 t
 ORIGIN

FEATURES
 source
 Query Match 1.3%; Score 39; DB 47; Length 520;
 Best Local Similarity 100.0%; Pred. No. 7.1e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3079 ccatacaaaaaa 523 bp mRNA EST 08-FEB-1999
 ||||||| 11 Subtracted Conidial Neurospora crassa cDNA clone SC5G11
 Db 479 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 517
 3' similar to ATP synthase subunit 4, mitochondrial precursor, mRNA
 sequence.

RESULT 44
 AI397624/C
 LOCUS
 DEFINITION
 AI397624 523 bp mRNA EST 08-FEB-1999
 NCSC5G117 Subtracted Conidial Neurospora crassa cDNA clone SC5G11
 3' similar to ATP synthase subunit 4, mitochondrial precursor, mRNA
 sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AI397624
AI397624.1 GI:4240709
EST.
Neurospora crassa
Neurospora crassa
Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae;
Neurospora,
1 (bases 1 to 523)

REFERENCE
AUTHORS

Nelson, M.A., Kang, S., Braun, E.L., Crawford, M.E., Dolan, P.L.,
Leonard, P.M., Mitchell, J., Armijo, A.M., Bean, L., Blueyes, E.,
Cushing, T., Errett, A., Fleharty, M., Gorman, M., Judson, K.,
Miller, R., Ortega, J., Pavlova, I., Pera, J., Todisco, S.,
Trujillo, R., Valentine, J., Wells, A., Werner-Washburne, M., Yazzie, S.,
and Natvig, D.O.
Expressed sequences from conidial, mycelial, and sexual stages of
Neurospora crassa
Fungal Genet. Biol. 21, 348-363 (1997)
97435549
On Apr 7, 1998 this sequence version replaced gi:3036493.
Department of Biology
University of New Mexico
Castetter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 3411
Fax: 505 277 0304
Email: ngp@biology.unm.edu.

TITLE

Expressed sequences from conidial, mycelial, and sexual stages of

JOURNAL
MEDLINE
COMMENT

Neurospora crassa
Fungal Genet. Biol. 21, 348-363 (1997)
97435549
On Apr 7, 1998 this sequence version replaced gi:3036493.
Department of Biology
University of New Mexico
Castetter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 3411
Fax: 505 277 0304
Email: ngp@biology.unm.edu.

FEATURES

source

1..523
/organism="Neurospora crassa"
/strain="74-OR23-IV A (FGSC 2489)"
/db_xref="taxon:5141"
/clone="SC5G11"
/clone_lib="Subtracted Conidial"
/sex="Mating Type A"
/tissue_type="Conidia"
/dev_stage="Germinating conidia"
/lab_host="E. coli"
/note="Vector: pBluescript SK (-); Site_1: EcoRI; Site_2:
XhoI; mRNA isolated from germinating conidia, grown in 1x
Vogel's, 2% sucrose for 4.5 hours. cDNA directionally
cloned into pBluescript SK(-) using the Uni-ZAP XR vector
system (Stratagene, La Jolla, CA). Previously identified
highly expressed clones were subtracted from this
library."
BASE COUNT 125 a 131 c 99 g 168 t
ORIGIN

Query Match 1.3%; Score 39; DB 74; Length 523;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 478 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 516

Search completed: September 10, 2000, 23:22:18
Job time: 6096 sec

BASE COUNT
ORIGIN

125 a 131 c 99 g 168 t

Query Match
Best Local Similarity
Matches

1.3%; Score 39; DB 39; Length 523;
100.0%; Pred. No. 7.1e-08;
39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079
Db 63

ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 25

RESULT 45
AW566146
LOCUS

AW566146 523 bp mRNA EST 10-MAR-2000
60062F05.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AW566146
AW566146.1 GI:7227505
EST.
Zea mays.
Zea mays
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
1 (bases 1 to 523)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University

REFERENCE
AUTHORS
TITLE

Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University

JOURNAL
COMMENT

Unpublished (1999)
On May 7, 1998 this sequence version replaced gi:3118896.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660062 row: F column: 05.

FEATURES
source

1..523
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XL0UR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."

BASE COUNT 105 a 156 c 131 g 130 t 1 others
ORIGIN

Query Match 1.3%; Score 39; DB 74; Length 523;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 478 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 516

Search completed: September 10, 2000, 23:22:18
Job time: 6096 sec

*****STN Columbus *****

FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000

=> file medline

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	TOTAL
FULL ESTIMATED COST	SESSION	0.15	0.15

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000

FILE LAST UPDATED: 22 SEP 2000 (20000922/UP). FILE COVERS 1960 TO DATE.

MEDLINE has been reloaded to reflect the annual MeSH changes made by the National Library of Medicine for 2000. Enter HELP RLOAD for details.

The OLDMEDLINE file segment now contains data from 1958 through 1965. Enter HELP CONTENT for details.

Left, right, and simultaneous left and right truncation are available in the Basic Index. See HELP SFIELDS for details.

THIS FILE CONTAINS CAS REGISTRY NUMBERS FOR EASY AND ACCURATE SUBSTANCE IDENTIFICATION.

=> s epithelial sodium channel/ab,bi

99491 EPTHELIAL/BI
251712 SODIUM/BI
75167 CHANNEL/BI
5419672 AB/FA
166 EPTHELIAL SODIUM CHANNEL/AB
(EPTHELIAL(W)SODIUM(W)CHANNEL)BI (L)
AB/FA)
99491 EPTHELIAL/BI
251712 SODIUM/BI
75167 CHANNEL/BI
216 EPTHELIAL SODIUM CHANNEL/BI
(EPTHELIAL(W)SODIUM(W)CHANNEL)BI
L1 216 EPTHELIAL SODIUM CHANNEL/AB,BI

=> s transgen?/ab,bi

28596 TRANSGEN?/BI
5419672 AB/FA
21527 TRANSGEN?/AB
(TRANSGEN?/BI (L) AB/FA)
28596 TRANSGEN?/BI

L2 28596 TRANSGEN?/AB,BI

=> s l1 and l2

L3 5 L1 AND L2

=> d l - bib ab

YOU HAVE REQUESTED DATA FROM 5 ANSWERS .
CONTINUE? Y(N)?

L3 ANSWER 1 OF 5 MEDLINE
AN 200023558 MEDLINE
DN 2023558
TI Dysfunction of epithelial sodium transport: from human to mouse.
AU Bonny O; Hummler E
CS Institut de Pharmacologie et de Toxicologie, Universite de Lausanne,
Lausanne, Switzerland.
SO KIDNEY INTERNATIONAL, (2000 Apr) 57 (4) 1313-8. Ref: 46
Journal code: KVB. ISSN: 0085-2538.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals
EM 200007
EW 20000702
AB The highly amiloride-sensitive ***epithelial***
sodium
channel (ENaC) is an apical membrane constituent of salt-absorbing epithelia. In the kidney, the functional relevance of ENaC expression has been well established. ENaC mediates the aldosterone-dependent sodium reabsorption in the distal nephron and is involved in the regulation of blood pressure. Mutations in genes encoding ENaC subunits are causative for two human inherited diseases: Liddle's syndrome, a severe form of hypertension associated with ENaC hyperfunction, and pseudohypoaldosteronism (PHA-1), a salt-wasting syndrome caused by decreased ENaC function. ***Transgenic*** mouse technologies provide a useful tool to study the role of ENaC in vivo. Different mouse lines have been established in which each of the ENaC subunits was affected. The phenotypes observed in these mice demonstrated that each subunit is essential for survival and for regulation of sodium

transport in kidney and colon. Moreover, the alpha subunit plays a specific role in the control of fluid absorption in the airways at birth.
Such mice can now be used to study the role of ENaC in various organs and can serve as models to understand the pathophysiology of these human diseases.

L3 ANSWER 2 OF 5 MEDLINE
AN 2000213442 MEDLINE
DN 20213442
TI Salt-sensitive hypertension in endothelin-B receptor-deficient rats.
AU Garpey C E; Ohuchi T; Williams S C; Richardson J A; Yanagisawa M
CS Howard Hughes Medical Institute, University of Texas Center, Dallas, Texas 75390, USA.
SO JOURNAL OF CLINICAL INVESTIGATION, (2000 Apr) 105 (7) 925-33.
Journal code: HS7. ISSN: 0021-9738.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Abbreviated Index Medicus Journals; Priority Journals; Cancer Journals
EM 200007
EW 20000704
AB The role of the endothelin-B receptor (ET(B)) in vascular homeostasis is controversial because the receptor has both pressor and depressor effects in vivo. Spotting lethal (sl) rats carry a naturally occurring deletion in the ET(B) gene that completely abrogates functional receptor expression. Rats homozygous for this mutation die shortly after birth due to congenital distal intestinal aganglionosis. Genetic rescue of ET(B)(sl/sl) rats from this developmental defect using a dopamine-hydroxylase (DBH)-ET(B) ***transgene*** results in ET(B)-deficient adult rats. On a sodium-deficient diet, DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats both exhibit a normal arterial blood pressure, but on a high-sodium diet, the former are severely hypertensive. We find no difference in plasma renin activity or plasma aldosterone concentration between salt-fed wild-type, DBH-ET(B);ET(B)(+/+) or DBH-ET(B);ET(B)(sl/sl) rats, and acute responses to intravenous L-NAME and indomethacin are similar between DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats. Irrespective of diet, DBH-ET(B);ET(B)(sl/sl) rats exhibit increased circulating ET-1, and,

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

on a high-sodium diet, they show increased but incomplete hypotensive responses to acute treatment an ET(A)-antagonist. Normal pressure is restored in salt-fed DBH-ET(B);ET(B)(s/s) rats when the ***epithelial*** ***sodium*** ***channel*** is blocked with amiloride. We conclude that DBH-ET(B);ET(B)(s/s) rats are a novel single-locus genetic model of severe salt-sensitive hypertension. Our results suggest that DBH-ET(B);ET(B)(s/s) rats are hypertensive because they lack the normal tonic inhibition of the renal ***epithelial*** ***sodium*** ***channel***.

L3 ANSWER 3 OF 5 MEDLINE
AN 1999345384 MEDLINE
DN 99345384
TI Implication of ENaC in salt-sensitive hypertension.
AU Hummler E
CS Institut de Pharmacologie et de Toxicologie, Université de Lausanne, Switzerland. ehummli@pop-server.unil.ch
SO JOURNAL OF STEROID BIOCHEMISTRY AND MOLECULAR BIOLOGY, (1999 Apr-Jun) 69 (1-6) 385-90. Ref: 49
Journal code: AX4. ISSN: 0960-0760.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals; Cancer Journals
EM 19991001
EW 19991001
AB Arterial blood pressure is critically dependent on sodium balance. The kidney is the key player in maintaining sodium homeostasis. Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC). Direct evidence that of ENaC participates in blood pressure regulation has come from the molecular analysis of two human genetic diseases, Liddle's syndrome and pseudohypoaldosteronism type 1 (PHA-1). Both, increased sodium reabsorption despite low aldosterone levels in Liddle's patients and decreased sodium reabsorption despite high aldosterone levels in PHA-1 patients, demonstrated that ENaC is an effector for aldosterone action. Gene-targeting and classical ***transgenic*** technology

enable the generation of mouse models for these diseases and the analysis of the involvement of the ***epithelial*** ***sodium*** ***channel*** (ENaC) in the progress of these diseases. A first mouse model using alphaENaC ***transgenic*** knockout mice [alphaENaC(-/-)Tg] mimicked several clinical features of PHA-1, like salt-wasting, metabolic acidosis, high aldosterone levels, growth retardation and increased early mortality. Such mouse models will be necessary in testing the involvement of genetic and/or environmental factors like salt-intake in hypertension.

L3 ANSWER 4 OF 5 MEDLINE
AN 1999170548 MEDLINE
DN 99170548
TI Genetic disorders of membrane transport. V. The ***epithelial*** ***sodium*** ***channel*** and its implication in human diseases.
AU Hummler E; Horisberger J D
CS Institut de Pharmacologie et de Toxicologie, Université de Lausanne, CH-1005 Lausanne, Switzerland.
SO AMERICAN JOURNAL OF PHYSIOLOGY, (1999 Mar) 276 (3 Pt 1) G567-71. Ref: 34
Journal code: 3U8. ISSN: 0002-9513.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals
EM 199906
EW 19990603
AB The epithelial Na⁺ channel (ENaC) controls the rate-limiting step in the process of transepithelial Na⁺ reabsorption in the distal nephron, the distal colon, and the airways. Hereditary salt-losing syndromes have been ascribed to loss of function mutations in the alpha-, beta-, or gamma-ENaC subunit genes, whereas gain of function mutations (located in the COOH terminus of the beta- or gamma-subunit) result in hypertension due to Na⁺ retention (Liddle's syndrome). In mice, gene-targeting experiments have shown that, in addition to the kidney salt-wasting phenotype, ENaC was essential for lung fluid clearance in newborn mice. Disruption of the

alpha-subunit resulted in a complete abolition of ENaC-mediated Na⁺ transport, whereas knockout of the beta- or gamma-subunit had only minor effects on fluid clearance in lung. Disruption of each of the three subunits resulted in a salt-wasting syndrome similar to that observed in humans.

L3 ANSWER 5 OF 5 MEDLINE
AN 97471032 MEDLINE
DN 97471032
TI A mouse model for the renal salt-wasting syndrome pseudohypoaldosteronism.
AU Hummler E; Barker P; Talbot C; Wang Q; Verdumo C; Grubb B; Gatz J; Burnier M; Horisberger J D; Beermann F; Boucher R; Rossier B C
CS Institut de Pharmacologie et de Toxicologie de l'Université, Rue du Bugnon 27, CH-1005 Lausanne, Switzerland.
ehummli@pop-server.unil.ch
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1997 Oct 14) 94 (21) 11710-5.
Journal code: PV3. ISSN: 0027-8424.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals; Cancer Journals
EM 199801
EW 19980104
AB Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the absorption of sodium through the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC) made of three homologous subunits (alpha, beta, and gamma). In human, autosomal recessive mutations of alpha, beta, or gammaENaC subunits cause pseudohypoaldosteronism type 1 (PHA-1), a renal salt-wasting syndrome characterized by severe hypovolemia, high plasma aldosterone, hyponatremia, life-threatening hyperkalemia, and metabolic acidosis. In the mouse, inactivation of alphaENaC results in failure to clear lung liquid at birth and in early neonatal death, preventing the observation of a PHA-1 renal phenotype. ***Transgenic*** alphaENaC driven by a cytomegalovirus promoter in mice [alphaENaC(-/-)Tg] rescued the perinatal lethal pulmonary phenotype and partially restored Na⁺ transport in renal, colonic, and pulmonary

THIS PAGE BLANK (USPTO)

epithelia. At days 5-9, however, alphaENaC(-/-)Tg mice showed clinical features of severe PHA-1 with metabolic acidosis, urinary salt-wasting, growth retardation, and 50% mortality. Adult alphaENaC(-/-)Tg survivors exhibited a compensated PHA-1 with normal acid/base and electrolyte values but 6-fold elevation of plasma aldosterone compared with wild-type littermate controls. We conclude that partial restoration of ENaC-mediated Na⁺ absorption in this ***transgenic*** mouse results in a mouse model for PHA-1.

=> file stnguide

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	TOTAL
	SESSION		
FULL ESTIMATED COST	1.85		2.00

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000
USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT
COPYRIGHT (C) 2000 AMERICAN CHEMICAL SOCIETY, JAPAN SCIENCE AND TECHNOLOGY CORPORATION, AND FACHINFORMATIONSZENTRUM KARLSRUHE

FILE CONTAINS CURRENT INFORMATION.
LAST RELOADED: Sep 15, 2000 (20000915/UP).

=> file medline embase biosis inpadoc caplus

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	TOTAL
	SESSION		
FULL ESTIMATED COST	0.00		2.00

FILE 'MEDLINE' ENTERED AT 14:28:05 ON 23 SEP 2000
FILE 'EMBASE' ENTERED AT 14:28:05 ON 23 SEP 2000
COPYRIGHT (C) 2000 Elsevier Science B.V. All rights reserved.
FILE 'BIOSIS' ENTERED AT 14:28:05 ON 23 SEP 2000
COPYRIGHT (C) 2000 BIOSIS(R)
FILE 'INPADOC' ENTERED AT 14:28:05 ON 23 SEP 2000
COPYRIGHT (C) 2000 European Patent Office, Vienna (EPO)
FILE 'CAPLUS' ENTERED AT 14:28:05 ON 23 SEP 2000
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE 'HELP USAGE/TERMS' FOR DETAILS.
COPYRIGHT (C) 2000 AMERICAN CHEMICAL SOCIETY (ACS)

=> d his

(FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000)

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000
L1 216 S EPITHELIAL SODIUM CHANNEL/AB,BI
L2 28596 S TRANSGEN7/AB,BI
L3 5 S L1 AND L2

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS' ENTERED AT 14:28:05 ON 23 SEP 2000

=> s i3

'AB' IS NOT A VALID FIELD CODE

L4 23 L3

=> dup rem l4

PROCESSING COMPLETED FOR L4

L5 10 DUP REM L4 (13 DUPLICATES REMOVED)

=> d l- bib ab

YOU HAVE REQUESTED DATA FROM 10 ANSWERS -
CONTINUE? Y/(N)/Y

L5 ANSWER 1 OF 10 CAPLUS COPYRIGHT 2000 ACS
AN 2000:401983 CAPLUS
DN 133:39099

TI Activated cation channel of the osteoblast as a mechanism of bone anabolism

IN Hruska, Keith
PA Barnes-Jewish Hospital, USA
SO PCT Int. Appl., 19 pp.

CODEN: PLOXDX2

DT Patent
LA English
FAN/CNT 1

PATENT NO.	KIND DATE	APPLICATION NO.
DATE		

PI WO 2000034458 A1 20000615 WO 1999-US28828
19991203

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU,

CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,

IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA,

MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI,

SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ,

BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE,

DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF,

CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
PRAI US 1998-110932 19981205

US 1998-111676 19981210

AB The present invention provides ***transgenic*** non-human mammals

which express stretch-activated cation channel ***transgene*** in

osteoblasts. Such pKBpA/alpha-ENaC animals are generated by transforming the mammals with a ***transgenic*** comprising

an

alpha-subunit ***epithelial*** ***sodium***

channel

(alpha-ENaC) cDNA or its variant inserted into a pKBpA gene and fused to

a promoter such as an osteocalcin promoter.

RE.CNT 4

RE

(1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541 CAPLUS

(2) Canessa, Nature 1993, V361, P467 CAPLUS

(3) Hummler, E, Journal of Steroid Biochemistry and Molecular Biology 1999,

V69(1-6), P385 CAPLUS

(4) Lingueglia, Federation of European Biochemical Societies 1993, V318(1), P95 CAPLUS

L5 ANSWER 2 OF 10 MEDLINE DUPLICATE 1
AN 2000223558 MEDLINE
DN 20223558

TI Dysfunction of epithelial sodium transport: from human to mouse.
AU Bonny O; Hummler E

CS Institut de Pharmacologie et de Toxicologie, Universite de Lausanne,

Lausanne, Switzerland

SO KIDNEY INTERNATIONAL, (2000 Apr) 57 (4) 1313-8. Ref. 46

Journal code: KVB. ISSN: 0085-2538.

CY United States

DT Journal; Article; (JOURNAL ARTICLE)

General Review; (REVIEW)

(REVIEW, TUTORIAL)

LA English

FS Priority Journals

EM 200007

EW 20000702

AB The highly amiloride-sensitive ***epithelial***

sodium

channel (ENaC) is an apical membrane constituent of cells of many

salt-absorbing epithelia. In the kidney, the functional relevance of ENaC

BEST AVAILABLE COPY

AVAILABLE COPY

THIS PAGE BLANK (USPTO)

expression has been well established. ENaC mediates the aldosterone-dependent sodium reabsorption in the distal nephron and is involved in the regulation of blood pressure. Mutations in genes encoding ENaC subunits are causative for two human inherited diseases: Liddle's syndrome, a severe form of hypertension associated with ENaC hyperfunction, and pseudohypoaldosteronism (PHA-1), a salt-wasting syndrome caused by decreased ENaC function. ***Transgenic*** mouse technologies provide a useful tool to study the role of ENaC in vivo. Different mouse lines have been established in which each of the ENaC subunits was affected. The phenotypes observed in these mice demonstrated that each subunit is essential for survival and for regulation of sodium transport in kidney and colon. Moreover, the alpha subunit plays a specific role in the control of fluid absorption in the airways at birth. Such mice can now be used to study the role of ENaC in various organs and can serve as models to understand the pathophysiology of these human diseases.

L5 ANSWER 3 OF 10 MEDLINE DUPLICATE 2
AN 2000213442 MEDLINE
DN 20213442
TI Salt-sensitive hypertension in endothelin-B receptor-deficient rats.
AU Gartezy C E; Ohuchi T; Williams S C; Richardson J A;
Yanagisawa M
CS Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA
SO JOURNAL OF CLINICAL INVESTIGATION, (2000 Apr) 105 (7) 925-33
Journal code: HS7. ISSN: 0021-9738.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Abridged Index Medicus Journals; Priority Journals; Cancer Journals
EM 200007
EW 20000704
AB The role of the endothelin-B receptor (ET(B)) in vascular homeostasis is controversial because the receptor has both pressor and depressor effects in vivo. Spotting lethal (sl) rats carry a naturally occurring deletion in the ET(B) gene that completely abrogates functional receptor expression. Rats homozygous for this mutation die shortly after birth due to

congenital distal intestinal aganglionosis. Genetic rescue of ET(B)(sl/sl) rats from this developmental defect using a dopamine-hydroxylase (DBH)-ET(B) ***transgene*** results in ET(B)-deficient adult rats. On a sodium-deficient diet, DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats both exhibit a normal arterial blood pressure, but on a high-sodium diet, the former are severely hypertensive. We find no difference in plasma renin activity or plasma aldosterone concentration between salt-fed wild-type, DBH-ET(B);ET(B)(+/+) or DBH-ET(B);ET(B)(sl/sl) rats, and acute responses to intravenous L-NAME and indomethacin are similar between DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats. Irrespective of diet, DBH-ET(B);ET(B)(sl/sl) rats exhibit increased circulating ET-1, and, on a high-sodium diet, they show increased but incomplete hypotensive responses to acute treatment an ET(A)-antagonist. Normal pressure is restored in salt-fed DBH-ET(B);ET(B)(sl/sl) rats when the ***epithelial*** ***sodium*** ***channel*** is blocked with amiloride. We conclude that DBH-ET(B);ET(B)(sl/sl) rats are a novel single-locus genetic model of severe salt-sensitive hypertension. Our results suggest that DBH-ET(B);ET(B)(sl/sl) rats are hypertensive because they lack the normal tonic inhibition of the renal ***epithelial*** ***sodium*** ***channel***.

L5 ANSWER 4 OF 10 MEDLINE DUPLICATE 3
AN 1999170548 MEDLINE
DN 99170548
TI Genetic disorders of membrane transport. V. The ***epithelial*** ***sodium*** ***channel*** and its implication in human diseases.
AU Hummler E; Horisberger J D
CS Institut de Pharmacologie et de Toxicologie, Universite de Lausanne, CH-1005 Lausanne, Switzerland.
SO AMERICAN JOURNAL OF PHYSIOLOGY, (1999 Mar) 276 (3 Pt 1) G567-71. Ref: 34
Journal code: 3U8. ISSN: 0002-9513.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals

EM 199906
EW 19990603
AB The epithelial Na⁺ channel (ENaC) controls the rate-limiting step in the process of transepithelial Na⁺ reabsorption in the distal nephron, the distal colon, and the airways. Hereditary salt-losing syndromes have been ascribed to loss of function mutations in the alpha-, beta-, or gamma-ENaC subunit genes, whereas gain of function mutations (located in the COOH terminus of the beta- or gamma-subunit) result in hypertension due to Na⁺ retention (Liddle's syndrome). In mice, gene-targeting experiments have shown that, in addition to the kidney salt-wasting phenotype, ENaC was essential for lung fluid clearance in newborn mice. Disruption of the alpha-subunit resulted in a complete abolition of ENaC-mediated Na⁺ transport, whereas knockout of the beta- or gamma-subunit had only minor effects on fluid clearance in lung. Disruption of each of the three subunits resulted in a salt-wasting syndrome similar to that observed in humans.

L5 ANSWER 5 OF 10 MEDLINE DUPLICATE 4
AN 1999345384 MEDLINE
DN 99345384
TI Implication of ENaC in salt-sensitive hypertension.
AU Hummler E
CS Institut de Pharmacologie et de Toxicologie, Universite de Lausanne, Switzerland. ehummler@pop-server.unil.ch
SO JOURNAL OF STEROID BIOCHEMISTRY AND MOLECULAR BIOLOGY, (1999 Apr-Jun) 69 (1-6) 385-90. Ref: 49
Journal code: AX4. ISSN: 0960-0760.
CY ENGLAND; United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals; Cancer Journals
EM 199910
EW 19991001
AB Arterial blood pressure is critically dependent on sodium balance. The kidney is the key player in maintaining sodium homeostasis. Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC). Direct evidence that

THIS PAGE BLANK (uspio)

dysfunction of ENaC participates in blood pressure regulation has come from the molecular analysis of two human genetic diseases, Liddle's syndrome and pseudohypoaldosteronism type 1 (PHA-1). Both, increased sodium reabsorption despite low aldosterone levels in Liddle's patients and decreased sodium reabsorption despite high aldosterone levels in PHA-1 patients, demonstrated that ENaC is an effector for aldosterone action. Gene-targeting and classical ***transgenic*** technology enable the generation of mouse models for these diseases and the analysis of the involvement of the ***epithelial*** ***sodium*** ***channel*** (ENaC) in the progress of these diseases. A first mouse model using alphaENaC ***transgenic*** knockout mice [alphaENaC(-/-)tg] mimicked several clinical features of PHA-1, like salt-wasting, metabolic acidosis, high aldosterone levels, growth retardation and increased early mortality. Such mouse models will be necessary in testing the involvement of genetic and/or environmental factors like salt-intake in hypertension.

L5 ANSWER 6 OF 10 EMBASE COPYRIGHT 2000 ELSEVIER SCI B.V.
AN 1998321597 EMBASE
TI ***Epithelial*** ***sodium*** ***channel*** and its implication in the control of blood pressure.
AU Hummler E.
CS Dr. E. Hummler, Institut Pharmacologie Toxicologie, 27 rue du Bugnon,
CH-1005 Lausanne, Switzerland. ehumm@pop-server.unil.ch
SO Kidney and Blood Pressure Research, (1998) 21/2-4 (253-255).
Refs: 14
ISSN: 1420-4096 CODEN: KBPRFC
CY Switzerland
DT Journal; Conference Article
FS 028 Urology and Nephrology
029 Clinical Biochemistry
LA English

L5 ANSWER 7 OF 10 MEDLINE DUPLICATE 5
AN 97471032 MEDLINE
DN 97471032
TI A mouse model for the renal salt-wasting syndrome pseudohypoaldosteronism.
AU Hummler E; Barker P; Talbot C; Wang Q; Verdumo C; Grubb B; Gatz J;
Burnier M; Horisberger J D; Beermann F; Boucher R; Rossier B C
CS Institut de Pharmacologie et de Toxicologie de l'Université, Rue du Bugnon

27, CH-1005 Lausanne, Switzerland. ehumm@pop-server.unil.ch
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1997 Oct 14) 94 (21) 11710-5.
Journal code: PV3 ISSN: 0027-8424.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals; Cancer Journals
EM 199801
EW 19980104
AB Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the absorption of sodium through the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC) made of three homologous subunits (alpha, beta, and gamma). In human, autosomal recessive mutations of alpha, beta, or gammaENaC subunits cause pseudohypoaldosteronism type 1 (PHA-1), a renal salt-wasting syndrome characterized by severe hypovolemia, high plasma aldosterone, hyponatremia, life-threatening hyperkalemia, and metabolic acidosis. In the mouse, inactivation of alphaENaC results in failure to clear fetal lung liquid at birth and in early neonatal death, preventing the observation of a PHA-1 renal phenotype. ***Transgenic*** alphaENaC driven by a cytomegalovirus promoter in alphaENaC(-/-) knockout mice [alphaENaC(-/-)tg] rescued the perinatal lethal pulmonary phenotype and partially restored Na+ transport in renal, colonic, and pulmonary epithelia. At days 5-9, however, alphaENaC(-/-)tg mice showed clinical features of severe PHA-1 with metabolic acidosis, urinary salt-wasting, growth retardation, and 50% mortality. Adult alphaENaC(-/-)tg survivors exhibited a compensated PHA-1 with normal acid/base and electrolyte values but 6-fold elevation of plasma aldosterone compared with wild-type littermate controls. We conclude that partial restoration of ENaC-mediated Na+ absorption in this ***transgenic*** mouse results in a mouse model for PHA-1.

L5 ANSWER 8 OF 10 BIOSIS COPYRIGHT 2000 BIOSIS AN 1997:8766 BIOSIS
DN PREV19979307969
TI Genetic rescue of alpha-ENaC knockout mouse: Establishment of

an animal model for pseudohypoaldosteronism (PHA 1).
AU Hummler, Edith (1); Barker, Pierre; Beermann, Friedrich; Verdumo, Chantal;
Gatz, John; Boucher, Richard; Rossier, Bernard C.
CS (1) Inst. Pharmacol. Toxicol., Lausanne Switzerland
SO Journal of the American Society of Nephrology, (1996) Vol. 7, No. 9, pp. 1281.
Meeting Info.: 29th Annual Meeting of the American Society of Nephrology
New Orleans, Louisiana, USA November 3-6, 1996
ISSN: 1046-6673.
DT Conference; Abstract; Conference
LA English

L5 ANSWER 9 OF 10 BIOSIS COPYRIGHT 2000 BIOSIS
AN 1995:242043 BIOSIS
DN PREV199598256343
TI Analysis of the ***epithelial*** ***sodium*** ***channel*** (ENaC) in ***transgenic*** mice.
AU Hummler, E. (1); Verdumo, C. (1); Beermann, F.; Rossier, B.
(1)
CS (1) Inst. Pharmacol. Toxicol., Lausanne Switzerland
SO Experimentia (Basel), (1995) Vol. 51, No. ABSTR., pp. A7.
Meeting Info.: 27th Annual Meeting of the Swiss Societies for Experimental Biology (USGEB/USSBE) Fribourg, Switzerland March 30-31, 1995
ISSN: 0014-4754.
DT Conference
LA English

L5 ANSWER 10 OF 10 BIOSIS COPYRIGHT 2000 BIOSIS
AN 1994:239206 BIOSIS
DN PREV199497252206
TI Analysis of the ***epithelial*** ***sodium*** ***channel*** (alpha-rENaC) in ***transgenic*** mice.
AU Hummler, E. (1); Verdumo, C. (1); Beermann, F.; Rossier, B.
(1)
CS (1) Inst. Pharmacol. Toxicol., Lausanne Switzerland
SO Experimentia (Basel), (1994) Vol. 50, No. ABSTR., pp. S40.
Meeting Info.: 26th Annual Meeting of the Swiss Societies for Experimental Biology (USGEB/USSBE) Bern, Switzerland March 17-18, 1994
ISSN: 0014-4754.
DT Conference
LA English

=> s 14 and osteoblast#ab/bi

'AB' IS NOT A VALID FIELD CODE
L6 1 LA AND OSTEOBLAST#AB,BI

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)



=> d bib ab

L6 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2000 ACS
 AN 2000-401983 CAPLUS
 DN 133-39099
 TI Activated cation channel of the ***osteoblast*** as a
 mechanism of
 bone anabolism
 IN Hruska, Keith
 PA Barnes-Jewish Hospital, USA
 SO PCT Int. Appl., 19 pp.
 CODEN: PIXXD2
 DT Patent
 LA English
 FAN/CNT 1
 PATENT NO. KIND DATE APPLICATION NO.
 DATE
 PI WO 2000034458 AI 20000615 WO 1999-US28828
 19991203
 W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH,
 CN, CR, CU,
 CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR,
 HU, ID, IL,
 IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU,
 LV, MA,
 MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU,
 SD, SE, SG, SI,
 SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,
 ZW, AM, AZ,
 BY, KG, KZ, MD, RU, TJ, TM
 RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT,
 BE, CH, CY, DE,
 DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF,
 BJ, CF,
 CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
 PRAI US 1998-110932 19981205
 US 1998-111676 19981210
 AB The present invention provides ***transgenic*** non-human
 mammals
 which express stretch-activated cation channel ***transgene***
 in
 osteoblasts. Such pKbPa/alpha-rENaC animals are
 generated by
 transforming the mammals with a ***transgenic*** comprising
 an
 alpha-subunit ***epithelial*** ***sodium***
 channel***
 (alpha-ENaC) cDNA or its variant inserted into a pKbPa gene
 and fused to
 a promoter such as an osteocalcin promoter.
 RE/CNT 4
 RE
 (1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541
 CAPLUS

(2) Canessa, Nature 1993, V361, P467 CAPLUS
 (3) Hummler, E, Journal of Steroid Biochemistry and Molecular
 Biology 1999,
 V69(1-6), P385 CAPLUS
 (4) Lingueglia, Federation of European Biochemical Societies 1993,
 V318(1), P95
 CAPLUS

=> d his

(FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000)

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000

L1 216 S EPIITHELIAL SODIUM CHANNEL/AB,BI

L2 28596 S TRANSGEN/AB,BI

L3 5 S L1 AND L2

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS'

ENTERED AT 14:28:05 ON 23

SEP 2000

L4 23 S L3

L5 10 DUP REM L4 (13 DUPLICATES REMOVED)

L6 1 S L4 AND OSTEOBLAST/AB,BI

=> s l4 and osteocalcin/ab,bi

'AB' IS NOT A VALID FIELD CODE

L7 1 L4 AND OSTEOCALCIN/AB,BI

=> d

L7 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2000 ACS

AN 2000-401983 CAPLUS

DN 133-39099

TI Activated cation channel of the osteoblast as a mechanism of bone

anabolism

IN Hruska, Keith

PA Barnes-Jewish Hospital, USA

SO PCT Int. Appl., 19 pp.

CODEN: PIXXD2

DT Patent

LA English

FAN/CNT 1

PATENT NO. KIND DATE APPLICATION NO.

DATE

PI WO 2000034458 AI 20000615 WO 1999-US28828

19991203

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH,

CN, CR, CU,

CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR,

HU, ID, IL,

IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,
 LV, MA,
 MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU,
 SD, SE, SG, SI,
 SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,
 ZW, AM, AZ,
 BY, KG, KZ, MD, RU, TJ, TM
 RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT,
 BE, CH, CY, DE,
 DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF,
 BJ, CF,
 CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
 PRAI US 1998-110932 19981205
 US 1998-111676 19981210
 RE/CNT 4
 RE
 (1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541
 CAPLUS
 (2) Canessa, Nature 1993, V361, P467 CAPLUS
 (3) Hummler, E, Journal of Steroid Biochemistry and Molecular
 Biology 1999,
 V69(1-6), P385 CAPLUS
 (4) Lingueglia, Federation of European Biochemical Societies 1993,
 V318(1), P95
 CAPLUS

US 1998-111676 19981210

RE/CNT 4

RE

(1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541

CAPLUS

(2) Canessa, Nature 1993, V361, P467 CAPLUS

(3) Hummler, E, Journal of Steroid Biochemistry and Molecular

Biology 1999,

V69(1-6), P385 CAPLUS

(4) Lingueglia, Federation of European Biochemical Societies 1993,

V318(1), P95

CAPLUS

=> e hruska keith/au

E1 1 HRUSKA KAREL PROF ING DRSC/AU

E2 3 HRUSKA KATHLEEN S/AU

E3 53 -> HRUSKA KEITH/AU

E4 119 HRUSKA KEITH A/AU

E5 2 HRUSKA KEITH A JR/AU

E6 1 HRUSKA KELTH A/AU

E7 4 HRUSKA KRUNICA/AU

E8 18 HRUSKA L/AU

E9 2 HRUSKA L A/AU

E10 4 HRUSKA L L/AU

E11 5 HRUSKA L S/AU

E12 2 HRUSKA L W/AU

=> s e3-e6

L8 175 ('HRUSKA KEITH/AU OR 'HRUSKA KEITH A/AU

OR 'HRUSKA KEITH A

JR/AU OR 'HRUSKA KELTH A/AU)

=> s l8 and l1

<---User Break--->

u

SEARCH ENDED BY USER

If this message appears repeatedly, please notify the Help Desk.

Enter 'HELP STN' for information on contacting the nearest STN

Help

Desk by telephone or via SEND in the STNMAIL file.

THIS PAGE BLANK (U0PT0)

=> s 18 and 11

COMMAND INTERRUPTED

'AB' IS NOT A VALID FIELD CODE

L9 5 L8 AND L1

To see the field codes for search terms in an L-number, enter

'DISPLAY

QUER Y* followed by the L-number at an arrow prompt. To see the field

codes for search terms in a saved query, enter 'ACTIVATE' and the query name, followed by /Q at an arrow prompt.

=> file medline embase biosis inpadoc caplus

COST IN U.S. DOLLARS SINCE FILE TOTAL

ENTRY SESSION

FULL ESTIMATED COST 50.73 52.73

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)

SINCE FILE TOTAL

ENTRY SESSION

CA SUBSCRIBER PRICE -1.11 -1.11

FILE 'MEDLINE' ENTERED AT 14:35:21 ON 23 SEP 2000

FILE 'EMBASE' ENTERED AT 14:35:21 ON 23 SEP 2000

COPYRIGHT (C) 2000 Elsevier Science B.V. All rights reserved.

FILE 'BIOSIS' ENTERED AT 14:35:21 ON 23 SEP 2000

COPYRIGHT (C) 2000 BIOSIS(R)

FILE 'INPADOC' ENTERED AT 14:35:21 ON 23 SEP 2000

COPYRIGHT (C) 2000 European Patent Office, Vienna (EPO)

FILE 'CAPLUS' ENTERED AT 14:35:21 ON 23 SEP 2000

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE 'HELP USAGETERMS' FOR DETAILS.

COPYRIGHT (C) 2000 AMERICAN CHEMICAL SOCIETY (ACS)

=> s 18 and 11

'AB' IS NOT A VALID FIELD CODE

L10 5 L8 AND L1

=> dup rem 110

PROCESSING COMPLETED FOR L10

L11 3 DUP REM L10 (2 DUPLICATES REMOVED)

=> d 1 - bib ab

YOU HAVE REQUESTED DATA FROM 3 ANSWERS -

CONTINUE? Y(N)?

L11 ANSWER 1 OF 3 CAPLUS COPYRIGHT 2000 ACS

AN 2000-401983 CAPLUS

DN 133-39099

TI Activated cation channel of the osteoblast as a mechanism of bone anabolism

IN ***Hruska, Keith***

PA Barnes-Jewish Hospital, USA

SO PCT Int. Appl., 19 pp.

CODEN: PXXXX2

DT Patent

LA English

FAN CNT 1

PATENT NO. KIND DATE APPLICATION NO.

DATE

PI WO 2000034458 A1 20000615 WO 1999-US28828

19991203

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH,

CN, CR, CU,

CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR,

HU, ID, IL,

IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,

LV, MA,

MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU,

SD, SE, SG, SI,

SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,

ZW, AM, AZ,

BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT,

BE, CH, CY, DE,

DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BE,

BI, CF,

CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRAI US 1998-110932 19981205

US 1998-111676 19981210

AB The present invention provides transgenic non-human mammals which express

stretch-activated cation channel transgene in osteoblasts. Such

pKbPa/alpha.-rENaC animals are generated by transforming the

mammals with

a transgenic comprising an alpha.-subunit ***epithelial***

sodium ***channel*** (alpha.-ENaC) cDNA or its

variant

inserted into a pKbPa gene and fused to a promoter such as an

osteocalcin

promoter.

RE CNT 4

RE

(1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541

CAPLUS

(2) Canessa, Nature 1993, V361, P467 CAPLUS

(3) Hummler, E. Journal of Steroid Biochemistry and Molecular

Biology 1999,

V69(1-6), P385 CAPLUS

(4) Lingueglia, Federation of European Biochemical Societies 1993,

V318(1), P95

CAPLUS

L11 ANSWER 2 OF 3 CAPLUS COPYRIGHT 2000 ACS

AN 1997:355199 CAPLUS

DN 126-341495

TI Reconstitution of stretch-activated cation channels by expression of the

alpha.-subunit of the ***epithelial*** ***sodium***

channel cloned from osteoblasts. [Erratum to document

cited in

CA126:197809]

AU Kizer, Neil; Guo, Xiao-Li; ***Hruska, Keith***

CS Renal Division, Barnes-Jewish Hospital, Washington University

Medical

Center, St. Louis, MO, 63110, USA

SO Proc. Natl. Acad. Sci. U. S. A. (1997), 94(8), 4233

CODEN: PNASAG; ISSN: 0027-8424

PB National Academy of Sciences

DT Journal

LA English

AB On pp. 1013 and 1017 the unit of conductance "picosiemens" was

incorrectly

abbreviated as "psec" and "ps" instead of "pS".

L11 ANSWER 3 OF 3 BIOSIS COPYRIGHT 2000 BIOSIS

DUPLICATE 1

AN 1997:127414 BIOSIS

DN PREV199799419227

TI Reconstitution of stretch-activated cation channels by expression of the

alpha.-subunit of the ***epithelial*** ***sodium***

channel

cloned from osteoblasts.

AU Kizer, Neil; Guo, Xiao-Li; ***Hruska, Keith (1)***

CS (1) Renal Div., Barnes-Jewish Hosp., Washington Univ. Med. Cent., 216

South Kingshighway Blvd., St. Louis, MO 63110 USA

SO Proceedings of the National Academy of Sciences of the United States of

America. (1997) Vol. 94, No. 3, pp. 1013-1018.

ISSN: 0027-8424.

DT Article

LA English

AB Osteoblasts respond to repetitive strain by activating

stretch-activated,

nonspecific cation channels (SA-CAT) and increasing matrix

protein

production. SA-CAT channels are thought to be responsible for mechano-transduction in osteoblasts, although the molecular

identity of

the SA-CAT channel has previously been unknown. We have

demonstrated that

both the UMR-106 osteoblast-like cell line and human osteoblasts

in

primary culture express the alpha-subunit of the ***epithelial***

sodium ***channel*** (alpha-ENaC). The ENaC

gene product is

closely related to a class of proteins that confer touch sensitivity to

Caenorhabditis elegans and are referred to as degenerins. A cDNA

THIS PAGE BLANK (USPTO)

clone was obtained of the entire coding region of rat alpha-ENaC (alpha-rENaC). Sequence analysis indicated that the osteoblast clone's sequence was identical to that originally cloned from rat colon. The alpha-rENaC cDNA was cloned into an expression plasmid and transfected into LM(TK-) cells, a null cell for SA-CAT activity. Stable transfectants expressed mRNA and the expected 74-kDa protein corresponding to alpha-rENaC. Reconstitution of alpha-rENaC resulted in the expression of a 24.2 +/- 1.0 psec channel (P-Na:P-K = 1.1 +/- 0.1). The channel is calcium permeable (P-Na:P-Ca = 1.4 +/- 0.1) and highly selective for cations over anions (P-Na:P-Cl mchgr 20). The channel is only active after negative pressure is applied to cell attached patches, cell swelling, or patch excision. These results represent the first heterologous expression of an SA-CAT channel in a mammalian cell system and provide evidence that the ENaC/degennin family of proteins are capable of mediating both transepithelial sodium transport and are involved in signal transduction by mechano-sensitive cells such as osteoblasts.

=> d his

(FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000)

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000

L1 216 S EPITHELIAL SODIUM CHANNEL/AB,BI
L2 28596 S TRANSGEN?/AB,BI
L3 5 S L1 AND L2

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS' ENTERED AT 14:28:05 ON 23 SEP 2000

L4 23 S L3
L5 10 DUP REM L4 (13 DUPLICATES REMOVED)
L6 1 S L4 AND OSTEOBLAST#/AB,BI
L7 1 S L4 AND OSTEOCALCIN/AB,BI
L8 E HRUSKA KEITH/AU
L8 175 S E3-E6
L9 5 S L8 AND L1

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS' ENTERED AT 14:35:21 ON 23 SEP 2000

L10 5 S L8 AND L1

L11 3 DUP REM L10 (2 DUPLICATES REMOVED)

=>

--Logging off of STN--

=>

Executing the logoff script...

=> LOG Y

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	SESSION	TOTAL
FULL ESTIMATED COST		17.57		70.30

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)
SINCE FILE TOTAL

CA SUBSCRIBER PRICE	ENTRY	SESSION
	-1.11	-2.22

STN INTERNATIONAL LOGOFF AT 14:36:22 ON 23 SEP 2000

THIS PAGE BLANK (U9PT9)